

Result No.	Score	Query Match	Length	DB	ID	Description
1	861	100.0	861	9	US-09-934-249-3	Sequence 3, Appl 1
2	861	100.0	864	16	US-10-295-027-127	Sequence 127, Appl 1
3	861	100.0	1321	9	US-09-934-249-1	Sequence 1, Appl 1
4	861	100.0	4839	15	US-10-241-230-119	Sequence 119, Appl 1
5	841	100.0	4839	15	US-10-265-909-84	Sequence 84, Appl 1
6	861	100.0	4839	15	US-10-265-909-85	Sequence 85, Appl 1
7	755-2	87.7	369	10	US-09-796-733-55	Sequence 55, Appl 1
8	754-2	87.6	1141	15	US-10-208-823-412	Sequence 412, Appl 1
9	754-2	87.6	1141	15	US-10-301-822-308	Sequence 208, Appl 1
10	754-2	87.6	1850	15	US-10-241-230-44	Sequence 44, Appl 1
11	754-2	87.6	4527	10	US-09-821-812-2	Sequence 2, Appl 1
12	753-6	87.5	1140	15	US-10-390-045-1	Sequence 1, Appl 1
13	752-8	87.4	759	15	US-10-390-045-2	Sequence 2, Appl 1
14	752-6	87.4	806	15	US-10-241-230-45	Sequence 45, Appl 1


```

; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/630,334
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: 09/606,565
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/606,317
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/665,666
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 09/677,751
; PRIOR FILING DATE: 2000-09-30
; NUMBER OF SEQ ID NOS: 162
; SEQ ID NO 55
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6)...(761)
US-09-796-753-55

Query Match
Best Local Similarity 99.6%; Score 755.2; DB 10; Length 969;
Matches 757; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 102 GGAGATCAGCGAGCTGGAGTTTCTTCAGATCATCATCGTGGTGGTGCATGATGATGAT 161
Db 2 GGAGATGGCGAGCTGGAGTTTCTTCAGATCATCATCGTGGTGGTGCATGATGATGAT 61

Qy 162 GGTGGTGGTGCATCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 221
Db 62 GGTGGTGGTGCATCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 121

Qy 222 CCGGCACAGCCAGCGCGGAGGAGAGAGATGCTCTCAGAGAGGATGCTGGTGGCC 281
Db 122 CCGGCACAGCCAGCGCGGAGGAGAGAGATGCTCTCAGAGAGGATGCTGGTGGCC 181

Qy 282 CTCGGAGACACAGTGTACGCAACCGAATCCAGAGCCGAGTTCAGCCCGGCTCG 341
Db 182 CTCGGAGACACAGTGTACGCAACCGAATCCAGAGCCGAGTTCAGCCCGGCTCG 241

Qy 342 GCCACACGCGCTGCGCGTGGCGCTTCGCCCGGAGCGCTTCACCGCTTCCA 401
Db 242 GCCACACGCGCTGCGCGTGGCGCTTCGCCCGGAGCGCTTCACCGCTTCCA 301

Qy 402 GCCACACTATCCGTACCTGCAGCAGCAGATCGACCTGCCACCCACCATCTTCGTGTGAGA 461
Db 302 GCCACACTATCCGTACCTGCAGCAGCAGATCGACCTGCCACCCACCATCTTCGTGTGAGA 361

Qy 462 CGGGGAGAGCCGCCACCTACAGGCGCTGCACCTGCACCTTCAGTTCGGGACCCCGAGCA 521
Db 362 CGGGGAGAGCCGCCACCTACAGGCGCTGCACCTGCACCTTCAGTTCGGGACCCCGAGCA 421

Qy 522 GCAGCTGGAACTGAACCGGAGTGGTGGCGGACCCCAACAGAACCATCTTCGACAG 581
Db 422 GCAGCTGGAACTGAACCGGAGTGGTGGCGGACCCCAACAGAACCATCTTCGACAG 481

Qy 582 TGACCTGATGATAGTGCAGGCTGGCGGCGCTTCGCCCGCCAGCAGTAACTTCGGGCT 641
Db 482 TGACCTGATGATAGTGCAGGCTGGCGGCGCTTCGCCCGCCAGCAGTAACTTCGGGCT 541

Qy 642 CAGCGCACAGTGTACGCGCAGCGCGATGAGAGGCGCGCGCCGACCATCTACAGCGA 701
Db 542 CAGCGCACAGTGTACGCGCAGCGCGATGAGAGGCGCGCGCCGACCATCTACAGCGA 601

Qy 702 GGTTCATCGGCCACTACCGGGGCTCTCTTCAGCAGCAGCAGAGTGGCGCGGCTC 761
Db 602 GGTTCATCGGCCACTACCGGGGCTCTCTTCAGCAGCAGCAGAGTGGCGCGGCTC 661

Qy 762 CTTCTCGAGGGGACCGGGCTCCACCAACACATCGCGCCCTAGAGAGCCAGCAT 821
Db 662 CTTCTCGAGGGGACCGGGCTCCACCAACACATCGCGCCCTAGAGAGCCAGCAT 721
```

RESULT 8

```

US-10-205-823-412
; Sequence 412, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Womsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; PRIOR FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 412
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-205-823-412
```

```

Query Match 87.6%; Score 754.2; DB 15; Length 1141;
Best Local Similarity 98.3%; Pred. No. 1.1e-189;
Matches 762; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 87 TTTGTTCCAGAGCATGGAGATCAGGAGCTGGAGTTTGTTCAGATCATCATCATCGTGT 146
Db 77 TCTCTGCAACACAGGCATGGCGAGCTGGAGTTTGTTCAGATCATCATCATCGTGT 136

Qy 147 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 206
Db 137 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 196

Qy 207 ACGGTCCTTCATCAGCGGCACAGCGGCGGAGAGAGAGATGCCCTGTCTCAGA 266
Db 197 ACGGTCCTTCATCAGCGGCACAGCGGCGGAGAGAGAGATGCCCTGTCTCAGA 256

Qy 267 AGGATGCTGTGGCCCTCGGAGAGACAGTGTACGAGCAACCGAATCCAGAGCCGAGGT 326
Db 257 AGGATGCTGTGGCCCTCGGAGAGACAGTGTACGAGCAACCGAATCCAGAGCCGAGGT 316

Qy 327 CTACGCCCGGCTCGGCCACCGACCGCTGGCGTGGCGCTTCGCCCGAGCGGAGCG 386
Db 317 CTACGCCCGGCTCGGCCACCGACCGCTGGCGTGGCGCTTCGCCCGAGCGGAGCG 376

Qy 387 CTTCCACCGCTTCAGAGCCCACTATCCGTACCTGCAGCAGAGATCGACCTGCCACCCAC 446
Db 377 CTTCCACCGCTTCAGAGCCCACTATCCGTACCTGCAGCAGAGATCGACCTGCCACCCAC 436
```

```
QY 447 CATCTCGTGTGACAGCGGAGAGAGCCGCCACCTTACCAGGGGCCCTGCACCCCTCCAGCT 506
Db 437 CATCTCGTGTGACAGCGGAGAGAGCCGCCACCTTACCAGGGGCCCTGCACCCCTCCAGCT 496
QY 507 TCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGGAGTGGTGGCGCGCACCCCAAACAG 566
Db 497 TCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGGAGTGGTGGCGCGCACCCCAAACAG 556
QY 567 AACCATCTTCGACAGTACCTGATAGTGTGACAGGCTGGCGCGGCCCTGCCGCCCCAG 626
Db 557 AACCATCTTCGACAGTACCTGATAGTGTGACAGGCTGGCGCGGCCCTGCCGCCCCAG 616
QY 627 CAGTAACTCGGGCATAGCCGACCTGCTACGAGCAGCGCGCGCGCATGAGAGGGCGGCC 686
Db 617 CAGTAACTCGGGCATAGCCGACCTGCTACGAGCAGCGCGCGCGCATGAGAGGGCGGCC 676
QY 687 GCCACCTACAGCAGGTGATCGGCCACTACCGGGGTCTCTCCAGCACCAGCAGAG 746
Db 677 GCCACCTACAGCAGGTGATCGGCCACTACCGGGGTCTCTCTCCAGCACCAGCAGAG 736
QY 747 CAGTGGCCCGCCCTCTTGTGAGGGGACCCGCTCCACACACACATCGCGGCCCT 806
Db 737 CAGTGGCCCGCCCTCTTGTGAGGGGACCCGCTCCACACACACATCGCGGCCCT 796
QY 807 AGAGAGCGGACCCATCTGGAGCAAGAGAGGATTAACAGAAAGGACACCCCTCTC 861
Db 797 AGAGAGCGGACCCATCTGGAGCAAGAGAGGATTAACAGAAAGGACACCCCTCTC 851

RESULT 9
US-10-301-822-208
; Sequence 208, Application US/10301822
; Publication No. US2003014840A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MEM01-029P2RM
; CURRENT APPLICATION NUMBER: US/10/301,822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 208
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (96)...(854)
US-10-301-822-208

Query Match 87.6%; Score 754.2; DB 15; Length 1141;
Best Local Similarity 98.3%; Pred No. 1.1e-189;
Matches 762; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 87 TTTGTTCCAGAGCATGGAGATCACGGAGCTGGAGTTTTCAGATCATCATCGTGGT 146
Db 77 TCTCTCGCAAAACAGGCAATGGCGAGCTGGAGTTTTCAGATCATCATCGTGGT 136
```

```
QY 147 GGTGATGATGTGTGATGTGTGGTGTATCACGTGCTGTGCTGAGCCACTACAAAGCTGTCTGC 206
Db 137 GGTGATGATGTGTGATGTGTGGTGTATCACGTGCTGTGCTGAGCCACTACAAAGCTGTCTGC 196
QY 207 AGGTCTCTTCATCAGCCGCGCACAGCGGCGGAGAGAGAAAGATGCCCTGTCTCAGA 266
Db 197 AGGTCTCTTCATCAGCCGCGCACAGCGGCGGAGAGAGAAAGATGCCCTGTCTCAGA 256
QY 267 AGGATGCTGTGGCCCTCGGAGAGCAGAGTGTCAAGGCAACGGAATCCAGAGCCGCGAGT 326
Db 257 AGGATGCTGTGGCCCTCGGAGAGCAGAGTGTCAAGGCAACGGAATCCAGAGCCGCGAGT 316
QY 327 CTACGCCCGGCTTGGGCCACCGACCGGCTTGGCCCTTGGCCCTTGGCCCTTGGCCCTTGGCCCT 386
Db 317 CTACGCCCGGCTTGGGCCACCGACCGGCTTGGCCCTTGGCCCTTGGCCCTTGGCCCTTGGCCCT 376
QY 387 CTTCCACCGCTTCCAGCCGCTTCCGCTACCTGCTGAGCAGCAGAGATCGACCTGCCACCCAC 446
Db 377 CTTCCACCGCTTCCAGCCGCTTCCGCTACCTGCTGAGCAGCAGAGATCGACCTGCCACCCAC 436
QY 447 CATCTCGTGTGACAGCGGAGAGAGCCGCCCTTACAGGGGCCCTTGCACCTTCCAGCT 506
Db 437 CATCTCGTGTGACAGCGGAGAGAGCCGCCCTTACAGGGGCCCTTGCACCTTCCAGCT 496
QY 507 TCGGACCCCGAGCAGCAGCTGGAACTGAACGGGGAGTGGTGGCGCGCATGGAGGGGCCGCC 566
Db 497 TCGGACCCCGAGCAGCAGCTGGAACTGAACGGGGAGTGGTGGCGCGCATGGAGGGGCCGCC 556
QY 567 AACCATCTTCGACAGTACCTGATAGTGTGACAGGCTGGCGCGGCCCTTGGCCGCCCTTGGCCGCCCT 626
Db 557 AACCATCTTCGACAGTACCTGATAGTGTGACAGGCTGGCGCGGCCCTTGGCCGCCCTTGGCCGCCCT 616
QY 627 CAGTAACTCGGGCATAGCCGACCTGCTACGAGCAGCGCGCGCGCATGGAGGGGCCGCC 686
Db 617 CAGTAACTCGGGCATAGCCGACCTGCTACGAGCAGCGCGCGCGCATGGAGGGGCCGCC 676
QY 687 GCCACCTACAGCAGGTGATCGGGCCACTACCGGGGTCTCTCTTCCAGACACCAAGCAGAG 746
Db 677 GCCACCTACAGCAGGTGATCGGGCCACTACCGGGGTCTCTCTTCCAGACACCAAGCAGAG 736
QY 747 CAGTGGCGGCCCTCTCTTGTGAGGGGACCCGCTCCACACACACATCGCGGCCCT 806
Db 737 CAGTGGCGGCCCTCTCTTGTGAGGGGACCCGCTCCACACACACATCGCGGCCCT 796
QY 807 AGAGAGCGGACCCATCTGGAGCAAGAGAGGATTAACAGAAAGGACACCCCTCTC 861
Db 797 AGAGAGCGGACCCATCTGGAGCAAGAGAGGATTAACAGAAAGGACACCCCTCTC 851

RESULT 10
US-10-241-220-44
; Sequence 44, Application US/10241220
; Publication No. US2003014840A1
; GENERAL INFORMATION:
; APPLICANT: Frantz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Phillips, Heidi
; APPLICANT: Polakis, Paul
; APPLICANT: Spencer, Susan
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF TUMOR
; FILE REFERENCE: P5010R1-US
; CURRENT APPLICATION NUMBER: US/10/241,220
; CURRENT FILING DATE: 2002-12-13
; NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 44
; LENGTH: 1850
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-241-220-44
```


TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED

Qy	822	CTGGAGCAAGAGAGAGGATAA	CAGAAAGGACACCCCTTC	861
Db	811	CTGGAGCAAGAGAGAGGATAA	CAGAAAGGACACCCCTTC	850

RESULT 13

US-10-390-045-2
; Sequence 2, Application US/10390045
; Publication NO. US20030170713A1
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, SHIV
; APPLICANT: MOUL, JUDD W.
; APPLICANT: XU, LINDA L.
; APPLICANT: SEGAWA, TAKEHIKO
; TITLE OF INVENTION: FOXYNUCLEOTIDE ARRAY
; TITLE OF INVENTION: FOXYNUCLEOTIDE ARRAY
; FILE REFERENCE: 04995.0057-00000
; CURRENT APPLICATION NUMBER: US/10/390,045
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US/09/769,482
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,045
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatenIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 759
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-390-045-2

Query Match 87.4%; Score 752.8; DB 15; Length 759;

Query Match	87.4%; Score 152.8; DB 13;	Length 133;
Best Local Similarity	99.7%; Pred. No. 2.5e-189;	
Matches 754: Conservative	0; Mismatches 2;	Indels 0; Gaps 0;

QY	106	ATCACGAGCTGGAGTTTGTTCAGATCATCATCGTGGTGATGATGGTCAATGGT	166
DB	1	ATGCGGAGCTGGAGTTTGTTCAGATCATCATCATCGTGGTGATGATGGTCAATGGT	60
QY	166	GTGGTGATCAGTGCCTGTGAGCCACTACAAGTGTCTGCACGGTGCTTTCATCAGCGGG	225
DB	61	GTGGTGATCAGTGCCTGTGAGCCACTACAAGTGTCTGCACGGTGCTTTCATCAGCGGG	120
QY	226	CACAGCCAGGGCGGAGGAGAGAAGATGCCCTGTCTCAGAAAGATGCCGTGTGCCCTCG	285
DB	121	CACAGCCAGGGCGGAGGAGAGAAGATGCCCTGTCTCAGAAAGATGCCGTGTGCCCTCG	180
QY	286	GAGAGCA CAGTGT CAGGCCAACGGAAATCC CAGAGCGCAGGTCTTACGCCCCGCCCTCGGCC	345
DB	181	GAGAGCA CAGTGT CAGGCCAACGGAAATCC CAGAGCGCAGGTCTTACGCCCCGCCCTCGGCC	240
QY	346	ACCGACCGCTGGCGGTGCCGCCCTTTCGCC CAGGGGAGCGCTTTCACCGCTTCACGCC	405
DB	241	ACCGACCGCTGGCGGTGCCGCCCTTTCGCC CAGGGGAGCGCTTTCACCGCTTCACGCC	300
QY	406	ACCTATCGGTACTG CAGCAGAGATCGA CCGTGCACCCACCACTCTCGCTGT CAGACGGG	465
DB	301	ACCTATCGGTACTG CAGCAGAGATCGA CCGTGCACCCACCACTCTCGCTGT CAGACGGG	360
QY	466	GAGAGCCCCACCCCTAC CAGGGGCCCTGTGC ACCCTCAGCTTCGGGACCCCGCAGCAGAG	525
DB	361	GAGAGCCCCCACCCCTAC CAGGGGCCCTGTGC ACCCTCAGCTTCGGGACCCCGCAGCAGAG	420
QY	526	CTGGAACTGAACCGGGAGTCCGTCCGGCGCACCCGCCAACAGAACCATCTTTCGACAGTGCAC	585
DB	421	CTGGAACTGAACCGGGAGTCCGTCCGGCGCACCCGCCAACAGAACCATCTTTCGACAGTGCAC	480
QY	586	CTGATGATAGTGCACGGCTGGGCGGGCCCCCTGCCCCCGCCACAGTAACTCGGGGCAATCAGC	645
DB	481	CTGATGATAGTGCACGGCTGGGCGGGCCCCCTGCCCCCGCCACAGTAACTCGGGGCAATCAGC	540

```
Query Match      37.5%; Score 753.6; DB 15; Length 1140;
Best Local Similarity 99.5%; Pred. NO. 1.6e-189;
Matches 756; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

[illegible]

222 CCGGCACAGCCAGGGCGGAGGAGAGAGATGCCCTGTCTCTCAGAGAGATGCTGTGGCC 281
211 CCGGCACAGCCAGGGCGGAGGAGAGATGCCCTGTCTCTCAGAGAGATGCTGTGGCC 270

282 CTGGAGAGCACAGTGTCTAGGCCAACGGAATCCAGAGCCGAGGTTACGCCCGCCCTCG 341
271 CTGGAGAGCACAGTGTCTAGGCCAACGGAATCCAGAGCCGAGGTTACGCCCGCCCTCG 330

342 GCCACCGACCGCTGGCGGTGCCGCCCTTCGCCCCAGCGGAGCGCTTCCACCGCTTCCA 401

402 GCCACCTATCCGTACCTGCAGACGAGATCGACCTGCCACCCACCATCTCGTGTGAGA 461
391 GCCACCTATCCGTACCTGCAGACGAGATCGACCTGCCACCCACCATCTCGTGTGAGA 450

462 CGGGGAGGAGCCCCACCCCTACCAAGGCCCCCTGCAACCTCTCCAGCTTCGGGACCCCGAGCA 521

451 CAGGCGGAGAGCCCCACCCCTACCAAGGCCCCCTGCAACCTCTCCAGCTTCGGGACCCCGAGCA 510

522 GCAGCTGGAAC TGAACCGGGAGTCGTGC GGCCACC CCACAAGAACATCTT CGACAG

570 GATGGTGTTAT GTATTGCCGGCCTGGGCTGG CCGCCCAATCATATCTGTTCTCTT

578 GATGGTGTTAT GTATTGCCGGCCTGGGCTGG CCGCCCAATCATATCTGTTCTCTT

582 TGRACCTGATGGATAGTCCAGGCTGGGGCGGGCCCTGCCCCCCCCCAGCAGTAACTCGGGCAT 641

642 CAGGGCCACGTGCTACGGCAGCGGGGGCGCATGGAGGGGGCGCCGCCACCTACAGCGA 701

691 CAGCGCCACGCGTCTATCGGCAGCGGGGGCATTCAGGGGGGCCCCCCTCTTCAAGCGA
702 GGTCATCGGCCACTACC CGGGGTCTCTCTTCAGCA CCAGCACAGCAGTGCGCGCCCTC 761

691 GG⁺CATCGGCGCAC⁺TACCGGGGTCCTCTCTCCAGGCATCCAGCAGAGCAGTGGGGCGCGCCCCC 751

762 CT⁺TGCTGGAGGGGACCGGGCTCCACCA⁺CACATCGCGCGCCTAGAGAGCGGAGCCAT 821

751 CTTGCTGGAGGGGACCCGGCTCCACCACACACACATCGCGCCCCCTAGAGAGCGCAGCCAT 811

QY 646 GCCAGTCTACGCGAGGGGCGCATGAGGGGCGCGCCACCTACAGGAGTTC 705
Db 541 GCCACGTCTACGCGAGGGGCGCATGAGGGGCGCGCCACCTACAGGAGTTC 600
QY 706 ATCGGCCACTACCGGGGTCTCTTCCAGCACAGAGAGAGTGGGGCGCCCTCTTG 765
Db 601 ATCGGCCACTACCGGGGTCTCTTCCAGCACAGAGAGAGTGGGGCGCCCTCTTG 660
QY 766 CTGAGGGGAGCCGGCTCCACACACACATCGGCCCTTAGAGGCGGAGCATCTGG 825
Db 661 CTGAGGGGAGCCGGCTCCACACACATCGGCCCTTAGAGGCGGAGCATCTGG 720
QY 826 AGCAAGAGAGAGGATAAAGAGGAGGACACCTCTC 861
Db 721 AGCAAGAGAGGATAAAGAGGAGGACACCTCTC 756

RESULT 14
US-10-241-220-45
; Sequence 45, Application US/10241220
; Publication No. US20030148408A1
; GENERAL INFORMATION:
; APPLICANT: Franz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Phillips, Heidi
; APPLICANT: Polakis, Paul
; APPLICANT: Spencer, Susan
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; FILE REFERENCE: P5010R1-US
; CURRENT APPLICATION NUMBER: US/10/241,220
; CURRENT FILING DATE: 2002-12-13
; NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 45
; LENGTH: 806
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-241-220-45

Query Match 87.4%; Score 752.6; DB 15; Length 806;
Best Local Similarity 98.2%; Pred. No. 2.8e-189;
Matches 761; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 87 TTGTTCAGAGATGGAGATCAGGAGCTGAGTTTGTTCAGATCATCATCGTGT 146
Db 14 TCTCTCGAAACAGGCAATGGCGAGCTGGAGTTTGTTCAGATCATCATCGTGT 73
QY 147 GGTGATGATGGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 206
Db 74 GGTGATGATGGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 133
QY 207 ACGGTCTTTCATAGCCGACACAGCAGGGGCGGAGAGAGAGATGCCCTGTCTCAGA 266
Db 134 ACGGTCTTTCATAGCCGACACAGCAGGGGCGGAGAGAGAGATGCCCTGTCTCAGA 193
QY 267 AGATGCTGTGGCCCTCGAGAGACAGTGTGTCAGGCAACGAAATCCCGAGCCGAGT 326
Db 194 AGATGCTGTGGCCCTCGAGAGACAGTGTGTCAGGCAACGAAATCCCGAGCCGAGT 253
QY 327 CTACGCCCCCGCTCGGACCAACCGACCGCTGCGCGCTTGGCCCGACCGAGCG 386
Db 254 CTACGCCCCCGCTCGGACCAACCGACCGCTGCGCGCTTGGCCCGACCGAGCG 313
QY 387 CTTCACCGCTTCAGCCGACACCTATCGGTACCTGACGACGACGATGACCTGCCAC 446
Db 314 CTTCACCGCTTCAGCCGACACCTATCGGTACCTGACGACGACGATGACCTGCCAC 373
QY 447 CATCTCGCTCAGAGCGGGAGGAGCCCGACCTACAGGGGCGCCCTGACACCTCGAGT 506

Db 374 CATCTCGCTGTACAGCGGGGAGGAGCCCGCCACCTTACAGGGGCGCCCTGACCTCCAGCT 433
QY 507 TCGGGACCCCGAGCAGAGCTGGAACTGAACCGGAGTTCGGCGCAGCCCGCAAAACAG 566
Db 434 TCGGGACCCCGAGCAGAGCTGGAACTGAACCGGAGTTCGGCGCAGCCCGCAAAACAG 493
QY 567 AACCATCTTCGACAGTACCTGATAGTAGTGCACAGGTTCGGCGGCGCCCTGCGCCCGCCAG 626
Db 494 AACCATCTTCGACAGTACCTGATAGTAGTGCACAGGTTCGGCGGCGCCCTGCGCCCGCCAG 553
QY 627 CAGTAACCTCGGCATCAGGCCACCTGCTACGGCAGCGCGCGCGCATGGAGGGGCGCGCC 686
Db 554 CAGTAACCTCGGCATCAGGCCACCTGCTACGGCAGCGCGCGCGCATGGAGGGGCGCGCC 613
QY 687 GCCACCTACAGCAGGTCATCGGCCATACCGGGGTCTCTTTCAGAGCAGCAGCAGAG 746
Db 614 GCCACCTACAGCAGGTCATCGGCCATACCGGGGTCTCTTTCAGAGCAGCAGCAGAG 673
QY 747 CAGTGGGCGCCCTCTCTGCTGGAGGGGAGCCCGGTCCACACACACACATCGCGCCCTC 806
Db 674 CAGTGGGCGCCCTCTCTGCTGGAGGGGAGCCCGGTCCACACACACACATCGCGCCCTC 733
QY 807 AGAGGCGCAGCCATCTGGAGCAAAAGAGAGATAAACAGAAAGAGACCCCTCTC 861
Db 734 AGAGGCGCAGCCATCTGGAGCAAAAGAGAGATAAACAGAAAGAGACCCCTCTC 788

RESULT 15
US-10-098-841-71
; Sequence 71, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungling
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
; FILE REFERENCE: 784CIP2
; CURRENT APPLICATION NUMBER: US/10/098,841
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 331
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 71
; LENGTH: 1066
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (154)...(867)
US-10-098-841-71

Query Match 86.2%; Score 742.4; DB 14; Length 1066;
Best Local Similarity 99.2%; Pred. No. 1.5e-186;
Matches 746; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 11, 2004, 21:06:02 ; Search time 390.646 Seconds

(without alignments)
9363.200 Million cell updates/sec

Title: US-09-934-249-3

Perfect score: 861

Sequence: 1 atgcaccgttgatgggggt.....aacagaaaggacacctctc 861

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq_29Jan04.*

1: Geneseqn1980s.*

2: Geneseqn1990s.*

3: Geneseqn2000s.*

4: Geneseqn2001as.*

5: Geneseqn2001bs.*

6: Geneseqn2002s.*

7: Geneseqn2003as.*

8: Geneseqn2003bs.*

9: Geneseqn2003cs.*

10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	861	100.0	1321	6	ABK12137 Human cDN
2	861	100.0	4839	7	ACC49552 Tumour-as
3	859.4	99.8	1383	9	ADC37326 Nuclear f
4	755.2	87.7	989	3	AAA75151
5	754.2	87.6	1140	6	ABK32120
6	754.2	87.6	1141	9	ADB75588
7	754.2	87.6	1850	7	ACC49536
8	753.6	87.5	969	3	AAA75164
9	753.6	87.5	969	3	AAA75163
10	753.6	87.5	969	3	AAA75165
11	753.6	87.5	1140	9	AAD50105
12	752.6	87.4	806	7	ACC49537
13	752.6	87.4	1061	3	AAA74229
14	752.6	87.4	1085	9	ADC37324
15	752.6	87.4	1334	7	ABZ36103
16	742.4	86.2	1066	4	AAI57868
17	712	82.7	1069	4	AAI59654
18	512.4	71.1	878	6	ABK12142
19	544.4	63.2	1383	6	ABK12142
20	462.2	53.7	1713	3	AAA75152
21	460.6	53.5	1713	3	AAA75167
22	460.6	53.5	1713	3	AAA75166
23	460.6	53.5	1713	3	AAA75168

C	24	445.6	51.8	693	6	ABK12143	Abk12143 Human MIV
	25	417.4	48.5	812	2	AAZ52964	Aaz52964 Human pro
C	26	417	48.4	474	7	ABZ84732	Abz84732 Toxicolog
	27	401.4	46.6	408	5	AAZ52983	Aaz52983 Novel hum
	28	400.6	46.5	673	6	ABT09178	Abt09178 Phase-1 R
	29	352.2	40.9	8494	5	AAZ77304	Aaz77304 DNA encod
	30	351.6	40.8	937	3	AAZ52507	Aaz52507 Human sec
	31	350	40.7	920	6	ABK34251	Abk34251 Human cDN
	32	350	40.7	2170	7	ADA52629	Ada52629 Human cod
	33	349	40.5	8065	2	AAV38335	Aav38335 Manic-dep
	34	349	40.5	8093	6	ABK12145	Abk12145 Human MIV
	35	341.8	39.7	475	6	ABK12144	Abk12144 Human MIV
C	36	333.8	38.8	8103	5	AAZ77312	Aaz77312 DNA encod
	37	320.4	37.2	1879	5	AAZ84503	Aaz84503 DNA encod
	38	284.2	33.0	1496	9	ADC37452	Adc37452 Nuclear f
	39	284.2	33.0	8440	5	AAZ77305	Aaz77305 DNA encod
	40	284.2	33.0	8440	6	ABK3477	Abk3477 Human cDN
	41	281	32.6	8011	2	AAV38336	Aav38336 Manic-dep
	42	252.6	29.3	8041	5	AAZ77310	Aaz77310 DNA encod
	43	232	26.9	254	3	AAA41265	Aaa41265 Human sec
	44	223.4	25.9	1428	5	AAZ92497	Aaz92497 DNA encod
	45	223.4	25.9	2942	5	AAZ77313	Aaz77313 DNA encod

ALIGNMENTS

RESULT 1
ABK12137
ID ABK12137 standard; cDNA; 1321 BP.
XX
AC ABK12137;
XX
DT 05-JUN-2002 (first entry)
XX
DE Human cDNA encoding Mechanically Induced Vascular Receptor 1, MIVR-1.
XX
KW Human; ss; Gene; MIVR-1; Mechanically Induced Vascular Receptor 1;
KW Cytostatic; cardiant; cerebroprotective; antiarteriosclerotic;
KW Cardiac cell; anti-apoptotic; vascular endothelial cell;
KW Cardiac hypertrophy; myocardial infarction; stroke; arteriosclerosis;
KW heart failure.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 413..1276
FT /*tag= a
FT /product= "MIVR-1"
FT /note= "This region is specifically claimed in claim 3"

XX WO200216416-A2.
XX 28-FEB-2002.
XX 21-AUG-2001; 2001WO-US026089.
XX 22-AUG-2000; 2000US-0227159P.
XX (BGHM) BEIGHAM & WOMENS HOSPITAL INC.
XX (PFIZ) PFIZER INC.
XX Lee RT, Landschulz KT, Kennedy SP, Thompson JP, Turi TG;
XX WPI; 2002-280912/32.
XX P-PSDB; AAU78231.
XX Novel nucleic acid molecule encoding Mechanically Induced Vascular
XX Receptor-1 polypeptide, useful for treating cardiovascular diseases.
XX Claim 2; Page 87-88; 105pp; English.
XX The invention relates to an isolated nucleic acid molecule encoding a

CC	Mechanically Induced Vascular Receptor (MIVR)-1 polypeptide having	
CC	cardiac cell anti-apoptotic activity and fragments of it provided they	
CC	are not identical to Genbank sequences A176141.1, A159439, NM_004338	
CC	and AQ17461. Also included are expression vectors, host cells, the MIVR-	
CC	1 polypeptide, MIVR-1 binding peptides, modulators of MIVR-1, contacting	
CC	a molecule having cardiac cell anti-apoptotic activity with a candidate	
CC	agent, where the molecule is a nucleic acid molecule comprising MIVR-1,	
CC	IFX-1, VDUP-1, BTG-2 and TIS-11d or its expression product, determining	
CC	if the anti-apoptotic activity is modulated and thereby identifying a	
CC	modulator. The cardiac cell anti-apoptotic molecules and nucleic acids	
CC	of the invention are useful for treating, diagnosing and monitoring	
CC	progression of such diseases and disorders as characterised by increased	
CC	apoptotic cell-death of vascular endothelial cells e.g. cardiac	
CC	hypertrophy, myocardial infarction, stroke, arteriosclerosis and heart	
CC	failure. The present sequence encodes human MIVR-1	
XX		
SQ	Sequence 1321 BP; 223 A; 493 C; 440 G; 165 T; 0 U; 0 Other;	
	Query Match 100.0%; Score 861; DB 6; Length 1321;	
	Best Local Similarity 100.0%; Pred. No. 1.8e-167;	
	Matches 861; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 ATGCACCGCTTGATGGGGGTCAACAGACACCGCGCGCCGCCGGGAGCCCAATGTC 60	
DB	413 ATGCACCGCTTGATGGGGGTCAACAGACACCGCGCGCGCGCGCGCCCAATGTC 472	
QY	61 TCCTGCAGCTGCRAACTGCRAACCGCTCTTTGTTCCAGAGCATGAGATCAGCGAGCTGGAG 120	
DB	473 TCCTGCAGCTGCRAACTGCRAACCGCTCTTTGTTCCAGAGCATGAGATCAGCGAGCTGGAG 532	
QY	121 TTTGTTCCAGATCATCATCTGTTGGTGATGATGTTGTTGTTGTTGTTGATCAGCTGC 180	
DB	533 TTTGTTCCAGATCATCATCTGTTGGTGATGATGTTGTTGTTGTTGATCAGCTGC 592	
QY	181 CTGCTGAGCCACTCAAGCTGTCTGCACGGTCTCTTATCATAGCCGCGACAGCCAGGGGCGG 240	
DB	593 CTGCTGAGCCACTCAAGCTGTCTGCACGGTCTCTTATCATAGCCGCGACAGCCAGGGGCGG 652	
QY	241 AGGAGAGAAGATGCCCTGTCTCTCAGAAGGATGCCCTGTGGCCCTTCGAGAGACACAGTGTCA 300	
DB	653 AGGAGAGAAGATGCCCTGTCTCTCAGAAGGATGCCCTGTGGCCCTTCGAGAGACACAGTGTCA 712	
QY	301 GGCAACGGAAATCCAGAGCCGAGGTCTACGCCCGCGCTCGGCCACCGACCGCTGGCC 360	
DB	713 GGCAACGGAAATCCAGAGCCGAGGTCTACGCCCGCGCTCGGCCACCGACCGCTGGCC 772	
QY	361 GTGCGCGCCTTCGCCCCAGCGGGAGCGGTCTCCACCGCTTCCAGCCCACTATCCGTACCTG 420	
DB	773 GTGCGCGCCTTCGCCCCAGCGGGAGCGGTCTCCACCGCTTCCAGCCCACTATCCGTACCTG 832	
QY	421 CAGCAGAGATGACCTGCCACCCACCAACATCTTCGCTGTTCAGACGGGAGAGCCCCACCC 480	
DB	833 CAGCAGAGATGACCTGCCACCCACCAACATCTTCGCTGTTCAGACGGGAGAGCCCCACCC 892	
QY	481 TACCAGGGCCCCCTGCACCCCTCCAGCTTCGGGACCCCGACGACGAGCTGGAACTGAACCGG 540	
DB	893 TACCAGGGCCCCCTGCACCCCTCCAGCTTCGGGACCCCGACGACGAGCTGGAACTGAACCGG 952	
QY	541 GAGTCGTTGCGCGCACCCCCAAAACAGAACCATCTTCGACAGTGACCTGATGGATAGTGCC 600	
DB	953 GAGTCGTTGCGCGCACCCCCAAAACAGAACCATCTTCGACAGTGATGGATAGTGATGCC 1012	
QY	601 AGGCTGGGCGGCCCTCTGCCCCCCCAGCAGTAACCTCGGGCATCAGCGCCACAGTGTACGGC 660	
DB	1013 AGGCTGGGCGGCCCTCTGCCCCCCCAGCAGTAACCTCGGGCATCAGCGCCACAGTGTACGGC 1072	
QY	661 AGCGCGGGCGCATGGAGGGCGCCGCCCACTTACAGCGAGGTCTATCGCCACTATCCG 720	
DB	1073 AGCGCGGGCGCATGGAGGGCGCCGCCCACTTACAGCGAGGTCTATCGCCACTATCCG 1132	
QY	721 GGGTCCTCTTCCAGCACACGACGACGAGTAGTGGCGCGCCCTCTTCTGCTGGAGGGGACCCGG 780	
DB	1133 GGGTCCTCTTCCAGCACACGACGACGAGTAGTGGCGCGCCCTCTTCTGCTGGAGGGGACCCGG 1192	

Qy	781	CTCCACACACACACATCGCGCCCTAGAGAGCGCAGCCACTCTGGAGCAAGAGAAGGAT	840
Db	1193	CTCCACACACACACATCGCGCCCTAGAGAGCGCAGCCACTCTGGAGCAAGAGAAGGAT	1252
Qy	841	AAACAGAAAGGACACCCCTCTC	861
Db	1253	AAACAGAAAGGACACCCCTCTC	1273
RESULT 2			
ACC49552	ID ACC49552 standard; cDNA; 4839 BP.		
XX	ACC49552;		
XX	01-JUL-2003 (first entry)		
XX	Tumour-associated antigenic target protein TAT379 SEQ ID NO:119.		
XX	Human; tumour-associated antigenic target; TAT; tumour; diagnosis;		
KW	cancer; gene; ss.		
KW			
XX	Homc sapiens.		
OS			
XX	WC2003024392-A2.		
FN			
XX	27-MAR-2003.		
PD			
XX	11-SEP-2002; 2002WO-US028959.		
FF			
XX	18-SEP-2001; 2001US-0323268P.		
PR	19-OCT-2001; 2001US-0339227P.		
PR	07-NOV-2001; 2001US-0336827P.		
PR	20-NOV-2001; 2001US-0331906P.		
PR	02-JAN-2002; 2002US-0345444P.		
PR	03-APR-2002; 2002US-0369724P.		
PR	19-AUG-2002; 2002US-0404809P.		
XX	(GETH) GENENTECH INC.		
FA			
PI	Frantz G, Hillan KJ, Phillips HS, Polakis P, Spencer SD;		
PI	Williams PM, Wu TD, Zhang Z;		
XX			
DR	WPI; 2003-354551/33.		
DR	P-PSDB; ABP97234.		
XX			
PT	New antibodies against tumor-associated antigenic target polypeptide,		
PT	useful for treating or diagnosing tumors or cancers in mammals, e.g.		
PT	prostate cancer, lung cancer, prostate adenocarcinomas or renal cell		
PT	carcinomas.		
XX			
PS	Claim 2; Fig 119; 285pp; English.		
XX			
CC	ACC49493 to ACC49552 encode the human tumour-associated antigenic target		
CC	(TAT) proteins given in ABP97175 to ABP97234. The present invention		
CC	describes an isolated antibody that binds to a polypeptide having at		
CC	least 80 % sequence identity to any of the 60 150-800 residue amino acid		
CC	sequences (S1), given in ABP97175 to ABP97234, comprising (S1), lacking		
CC	its associated signal peptide, encoded by any of the 60 2000-3000 base		
CC	pair sequences (S2), given in ACC49493 to ACC49552. TAT proteins have		
CC	cytostatic activity. The antibody can be used for treating or diagnosing		
CC	tumours or cancers in mammals, e.g. prostate cancer, lung cancer, breast		
CC	cancer, colon cancer, ovarian cancer, prostate adenocarcinomas, renal		
CC	cell carcinomas, or thyroid cancer		
XX			
SQ	Sequence 4839 BP; 1103 A; 1313 C; 1290 G; 1133 T; 0 U; 0 Other;		
Query Match 100.0%; Score 861; DB 7; Length 4839;			
Best Local Similarity 100.0%; Pred.No.2.le-167;			
Matches 861; Conservative 0; Mismatches 0; Indels 0; Gaps 0			

Db 741 CAGCAGGATCGACCTCGCGCCACCACATCTCGCTGTTCAGACGGGGAGGACCCGCCACCC 800
Qy 481 TACCAGGCGCCCTGCACCTCTCAGCTTCGGGACCCCGAGCAGCAGCTGGACTGAACGG 540
Db 801 TACCAGGCGCCCTGCACCTCTCAGCTTCGGGACCCCGAGCAGCAGCTGGACTGAACGG 860
Qy 541 GAGTCGGTGGCGGACCCGCCAAACAGAACCATCTTCGACAGTGCACCTGATGATGATGCC 600
Db 861 GAGTCGGTGGCGGACCCGCCAAACAGAACCATCTTCGACAGTGCACCTGATGATGATGCC 920
Qy 601 AGGCTGGGGCGCCCTCGCGCCCGCCAGCAGTAACTCGGGCATCAGCGCCACCTGCTAGCGC 660
Db 921 AGGCTGGGGCGCCCTCGCGCCCGCCAGCAGTAACTCGGGCATCAGCGCCACCTGCTAGCGC 980
Qy 661 AGCGCGGGCGCATGAGGGGCGCGCCCGCCACCTACAGCAGAGTTCATCGGCCCATACCCG 720
Db 981 AGCGCGGGCGCATGAGGGGCGCGCGCCCGCCACCTACAGCAGAGTTCATCGGCCCATACCCG 1040
Qy 721 GGGTCTCTTCCACACACAGCAGCAGTGGGCGCCCTCTGCTGGAGGGACCCGG 780
Db 1041 GGGTCTCTTCCACACACAGCAGCAGTGGGCGCCCTCTGCTGGAGGGACCCGG 1100
Qy 781 CTCACCCACACACATCGCGCCCTAGAGAGCGCAGCCATCTGGAGCAAGAGAGGAT 840
Db 1101 CTCACCCACACACATCGCGCCCTAGAGAGCGCAGCCATCTGGAGCAAGAGAGGAT 1160
Qy 841 AAACAGAAAGACACCTCTC 861
Db 1161 AAACAGAAAGACACCTCTC 1181

RESULT 4

ID AAA75151 standard; cDNA; 969 BP.
XX
AC AAA75151;
XX
DT 15-JAN-2001 (first entry)
XX
XX cDNA encoding a human TANGO 261 polypeptide.
XX
XW TANGO 266; TANGO 216; TANGO 261; TANGO 262; TANGO 267;
KW cellular proliferation; cellular differentiation; cellular adhesion;
KW von Willebrand factor-associated disorder; cell trafficking; cancer;
KW hematopoietic associated disease; atelectasis; pulmonary congestion;
KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;
KW intestinal disorder; spleen associated disease; renal disorder;
KW cardiovascular disorder; ischemic heart disease; hydrocephalus;
KW brain herniation; iatrogenic disease; inflammation; meningitis;
KW Alzheimer's disease; cerebral toxoplasmosis; Parkinson's disease;
KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.

OS Homo sapiens.
XX
XX
XX
FH Key Location/Qualifiers
FT CDS 6..764
FT /*tag= a
FT /product= "TANGO 261"
FT sig_peptide 6..89
FT /*tag= b
FT mat_peptide 90..764
FT /*tag= c
XX
FN WO200052022-A1.
XX
XX
PD 08-SEP-2000.
XX
XX 01-MAR-2000; 2000WO-US0005226.
XX
XX 01-MAR-1999; 99US-0122458P.
XX
XX (MILL-) MILLENNIUM PHARM INC.

PI Barnes TM, Holtzman DA, Sharp JD, Fraser CC;
XX WPI: 2000-579269/54.
DR P-PSDB; AAB18449.
XX
PT Novel human and murine secreted proteins designated TANGO 216, 261, 262,
PT 266 and 267 useful as modulating agents of cellular processes, e.g. for
PT treating cancer.
XX
PS Claim 2; Fig 5; 175pp; English.
XX
CC The present sequence encodes a human TANGO 261 polypeptide. The
CC specification also describes TANGO 266, TANGO 216, TANGO 262, and TANGO
CC 267. The TANGO polypeptides can be used to modulate cellular
CC proliferation, modulate cellular differentiation and/or modulate cellular
CC adhesion. The proteins can be used to treat any von Willebrand factor-
CC associated disorder, regulate extracellular matrix structuring, cellular
CC adhesion, and cell trafficking and/or migration, modulate cellular
CC interactions, modulate cell adhesion in proliferative disorders, such as
CC cancer, modulate the proliferation, differentiation, and/or function of
CC cells that appear in the bone marrow, and leukocytes, treat bone marrow,
CC blood and hematopoietic associated diseases and disorders, atelectasis,
CC pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial
CC asthma and bronchiectasis, intestinal disorders, spleen associated
CC diseases, modulate renal disorders, treat cardiovascular disorders such
CC as ischemic heart disease, modulate the proliferation, differentiation,
CC and/or function of bone and cartilage cells and to treat bone and/or
CC cartilage associated diseases or disorder. They may also be used to treat
CC disorders associated with the ovaries, cerebral oedema, hydrocephalus,
CC brain herniations, iatrogenic disease, inflammations, bacterial and viral
CC meningitis, Alzheimer's Disease, cerebral toxoplasmosis, Parkinson's
CC disease, multiple sclerosis, brain cancers, hydrocephalus and
CC encephalitis, and treat hepatic disorders
XX
SQ Sequence 969 BP; 211 A; 316 C; 294 G; 148 T; 0 U; 0 Other;

Query Match 87.7%; Score 755.2; DB 3; Length 969;
Best Local Similarity 99.6%; Pred. No. 9e-146;
Matches 757; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 102 GGAGATCAGGAGCTGGAGTTGTTTCAGATCATCATCTGCTGGTGTGATGATGATGAT 161
Db 2 GGAGATGCGGGAGCTGGAGTTGTTTCAGATCATCATCTGCTGGTGTGATGATGATGAT 61
Qy 162 GGTGGTGGTGTGATCAGCTGCTGCTGAGCCACTACAGCTGTCTGACGGTCTTATCAG 221
Db 62 GGTGGTGGTGTGATCAGCTGCTGCTGAGCCACTACAGCTGTCTGACGGTCTTATCAG 121
Qy 222 CCGGCACAGCCAGCGGGCGGAGAGATGCCCTGCTCTCAGAGAGGATGCTGTGGCC 281
Db 122 CCGGCACAGCCAGCGGGCGGAGAGATGCCCTGCTCTCAGAGAGGATGCTGTGGCC 181
Qy 282 CTGGAGAGCAGAGTGTTCAGGCAACGGAATCCAGAGCCGCGAGGTCTACGCCCGCTCG 341
Db 182 CTGGAGAGCAGAGTGTTCAGGCAACGGAATCCAGAGCCGCGAGGTCTACGCCCGCTCG 241
Qy 342 GCCACCGACCGCTGCGCGCTGCGCGCTTCCGCCAGCGGGAGCGCTTCCACCGCTTCCA 401
Db 242 GCCACCGACCGCTGCGCGCTGCGCGCTTCCGCCAGCGGGAGCGCTTCCACCGCTTCCA 301
Qy 402 GCCACCGCTATCCGTACCTGCGAGCAGAGATCCAGCTGCCACCCACCATCTCGTGTGAGA 461
Db 302 GCCACCGCTATCCGTACCTGCGAGCAGAGATCCAGCTGCCACCCACCATCTCGTGTGAGA 361
Qy 462 CGGGAGAGAGCCCGCCACCTTACAGGGCGCTGTCACCTCTCAGCTTGGGAGCCCGAGCA 521
Db 362 CGGGAGAGAGCCCGCCACCTTACAGGGCGCTGTCACCTCTCAGCTTGGGAGCCCGAGCA 421
Qy 522 GCAGCTGGAATGAACCGGGAGTGGTGGCGGACACCCCAACAGAACCATCTTTCGACAG 581
Db 422 GCAGCTGGAATGAACCGGGAGTGGTGGCGGACACCCCAACAGAACCATCTTTCGACAG 481
Qy 582 TGACCTGATGATAGTGCAGGCTGGGCGGCCCTGCGCCCGCCCGAGGATTAACCTCGGGCAT 641

AC ADB75588;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Prostate cancer marker cDNA.
 XX
 KW Prostate; cancer; cytostatic; gene therapy; marker; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO2003009814-A2.
 XX
 PD 06-FEB-2003.
 XX
 PF 25-JUL-2002; 2002WO-US023913.
 XX
 PR 25-JUL-2001; 2001US-0307982P.
 XX
 PR 22-AUG-2001; 2001US-0314356P.
 XX
 PR 25-SEP-2001; 2001US-0325020P.
 XX
 PR 12-DEC-2001; 2001US-0341745P.
 XX
 PR 05-MAR-2002; 2002US-0362158P.
 XX
 (MILL-) MILLENNIUM PHARM INC.
 PA
 XX Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbacheva B;
 PI Hoersh S, Kamatkar S, Woney AM, Glatt K, Zhao X, Anderson D;
 XX
 DR WPI; 2003-248033/24.
 XX
 XX New nucleic acid molecule, useful for diagnosing or treating prostate
 PT cancer.
 XX
 PS Disclosure; SEQ ID NO 412; 99pp; English.
 XX
 CC The invention relates to newly discovered cancer markers associated with
 CC the cancerous state of prostate cells. Also disclosed is a method of
 CC assessing whether a patient is afflicted with prostate cancer. The method
 CC of the invention involves assessing whether a patient is afflicted with
 CC prostate cancer by comparing the level of expression of a marker in a
 CC patient sample and the normal level of expression of the marker in a
 CC control non-prostate cancer sample, where a significant increase in the
 CC level of expression of the marker in the patient sample and the normal
 CC level indicates that the patient is afflicted with prostate cancer.
 CC Nucleic acids of the invention are useful for diagnosing or treating
 CC prostate cancer, and may be useful in gene therapy. Sequences given in
 CC ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1141 BP; 271 A; 350 C; 336 G; 184 T; 0 U; 0 Other;
 Query Match 87.6%; Score 754.2; DB 9; Length 1141;
 Best Local Similarity 98.3%; Pred. No. 1.5e-145;
 Matches 762; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
 QY 87 TTGTTCAGAGTGGAGATCAGCGAGTGGAGTTGTTCAGATCATCATCGTGGT 146
 DB 77 TCTCCGCGAAACCGGCAATGGCGAGTGGAGTTGTTCAGATCATCATCGTGGT 136
 QY 147 GGTGATGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 206
 DB 137 GGTGATGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 196
 QY 207 ACGTCTCTTCATCAGCCGCGACAGCCAGGGGGGAGAGAGATGGCTGCTCTCAGA 266
 DB 197 ACGTCTCTTCATCAGCCGCGACAGCCAGGGGGGAGAGAGATGGCTGCTCTCAGA 256
 QY 267 AGGATCCCTGTGGCCCTCGAGAGACAGTGTGAGGCAACGAAATCCAGAGCCCGAGGT 326
 DB 257 AGGATCCCTGTGGCCCTCGAGAGACAGTGTGAGGCAACGAAATCCAGAGCCCGAGGT 316
 QY 327 CTACGCCCGCCCTTCGCGCCACCGACCGCCCTGGCGGTGGCCCTTCGCCCGGAGCG 386

DB 317 CTACGCCCGCCCTTCGCGCCACCGACCGGCTGGCGGTGGCGGCTTCGCCCGGAGGCG 376
 QY 387 CTTTCCACCGCTTCCAGAGCCACCTATCTCGTACCTGAGCACGAGATCGACCTGCCACCCAC 446
 DB 377 CTTTCCACCGCTTCCAGAGCCACCTATCTCGTACCTGAGCACGAGATCGACCTGCCACCCAC 436
 QY 447 CATCTCGTGTTCAGACGGGGAGAGCCCGCCACCTTACAGGGCCCTTCGACCTCCAGCT 506
 DB 437 CATCTCGTGTTCAGACGGGGAGAGCCCGCCACCTTACAGGGCCCTTCGACCTCCAGCT 496
 QY 507 TCGGGACCCCGAGCAGCAGCTGGAATGAAACCGGAGTGGGTGGCGGACCCCGCCAAAACAG 566
 DB 497 TCGGGACCCCGAGCAGCAGCTGGAATGAAACCGGAGTGGGTGGCGGACCCCGCCAAAACAG 556
 QY 567 AACCATCTTCGACAGTGCATGATGATAGTGCAGGCTGGCGGCGCTTCGCCCGCCAG 626
 DB 557 AACCATCTTCGACAGTGCATGATGATAGTGCAGGCTGGCGGCGCTTCGCCCGCCAG 616
 QY 627 CAGTAACTCGGGCATCAGCGCCACCTGCTACCGGAGCGGGCGCATGGAGGGGCGCCGCC 686
 DB 617 CAGTAACTCGGGCATCAGCGCCACCTGCTACCGGAGCGGGCGCATGGAGGGGCGCCGCC 676
 QY 687 GCCACCTACAGCGAGGTATCGGCCACTACCGGGGCTCTCTTCAGACACACAGAGAG 746
 DB 677 GCCACCTACAGCGAGGTATCGGCCACTACCGGGGCTCTCTTCAGACACACAGAGAG 736
 QY 747 CAGTGGCGCGCCCTCTCTTCGAGGAGGACCGGCTCCACACACACACATCGCGCCCT 806
 DB 737 CAGTGGCGCGCCCTCTCTTCGAGGAGGACCGGCTCCACACACACACATCGCGCCCT 796
 QY 807 AGAGAGCGCAGCATCTCGAGCAAGAGAGATATAAAGAGAGAGAGAGAGAGAGAGAGAG 861
 DB 797 AGAGAGCGCAGCATCTCGAGCAAGAGAGATATAAAGAGAGAGAGAGAGAGAGAGAGAG 851
 RESULT 7
 ACC49536
 ID ACC49536 standard; cDNA; 1850 BP.
 XX
 AC ACC49536;
 XX
 DT 01-JUL-2003 (first entry)
 XX
 DE Tumour-associated antigenic target protein TAT180 cDNA SEQ ID NO:44.
 XX
 KW Human; tumour-associated antigenic target; TAT; tumour; diagnosis;
 KW cancer; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO2003024392-A2.
 XX
 PD 27-MAR-2003.
 XX
 PF 11-SEP-2002; 2002WO-US028859.
 XX
 PR 18-SEP-2001; 2001US-0323268P.
 XX
 PR 19-OCT-2001; 2001US-0339227P.
 XX
 PR 07-NOV-2001; 2001US-0336827P.
 XX
 PR 20-NOV-2001; 2001US-0331906P.
 XX
 PR 02-JAN-2002; 2002US-0345444P.
 XX
 PR 03-APR-2002; 2002US-0369724P.
 XX
 PR 19-AUG-2002; 2002US-0404809P.
 XX
 (GETH) GENENTECH INC.
 PA
 XX Frantz G, Hillan KJ, Phillips HS, Polakis P, Spencer SD;
 PI Williams PM, Wu TD, Zhang Z;
 XX
 DR WPI; 2003-354551/33.
 DR P-PSDB; ABP97218.
 XX

Db 2 GGAGATGGCGGACCTGGAGTTTGTTCAGATCATCATCGTGTGTGATGATGAT 61
Qy 162 GTGGTGGTGTGATCAGTGGCTGTGAGCCACTAAGAGCTGTCTGACGGTCTTTCATCAG 221
Db 62 GTGGTGGTGTGATCAGTGGCTGTGAGCCACTAAGAGCTGTCTGACGGTCTTTCATCAG 121
Qy 222 CCGGCACAGCCAGGGCGGAGAGAGAGTGCCTGTCTCAGAGGATGCTGTGGCC 281
Db 122 CCGGCACAGCCAGGGCGGAGAGAGAGTGCCTGTCTCAGAGGATGCTGTGGCC 181
Qy 282 CTCGGAGAGCACAGTGTTCAGGCAACGGAATCCAGAGCCGACGCTCTACGCCCGCTCG 341
Db 182 CTCGGAGAGCACAGTGTTCAGGCAACGGAATCCAGAGCCGACGCTCTACGCCCGCTCG 241
Qy 342 GCCCACCAGCGCTGCGCTGGCGCTTTCGCGCCAGCGGAGCGCTTCACCGCTTCCA 401
Db 242 GCCCACCAGCGCTGCGCTGGCGCTTTCGCGCCAGCGGAGCGCTTCACCGCTTCCA 301
Qy 402 GCCCACCATTCCTGACGACGAGATCGACCTGCCACCCACCACTCTCGTGTGAGA 461
Db 302 GCCCACCATTCCTGACGACGAGATCGACCTGCCCGCCACCACTCTCGTGTGAGA 361
Qy 462 CGGGAGAGCCCGCCACCTTACAGGCGCTGTGACCTTCAGCTTCGGGACCCCGAGCA 521
Db 362 CGGGAGAGCCCGCCACCTTACAGGCGCTGTGACCTTCAGCTTCGGGACCCCGAGCA 421
Qy 522 GCAGCTGGAATCTGAACCGGAGTGGTGGCGGACCCGCAACAGAACCACTTTCGACAG 581
Db 422 GCAGCTGGAATCTGAACCGGAGTGGTGGCGGACCCGCAACAGAACCACTTTCGACAG 481
Qy 582 TGACCTGATGATAGTGCAGGCTGGCGGCGCTGTGCGCCCGCCAGAGTAACCTCGGCGAT 641
Db 482 TGACCTGATGATAGTGCAGGCTGGCGGCGCTGTGCGCCCGCCAGAGTAACCTCGGCGAT 541
Qy 642 CAGCGCCAGCTGTACGCGCAGCGCGGCGCATGAGGGGCGCGCCGACCTTACAGCGA 701
Db 542 CAGCGCCAGCTGTACGCGCAGCGCGGCGCATGAGGGGCGCGCCGACCTTACAGCGA 601
Qy 702 GGTGATCGGCATACCGCGGCTCTCTTCAGCAGCAGAGAGTGGCGCGCCCTC 761
Db 602 GGTGATCGGCATACCGCGGCTCTCTTCAGCAGCAGAGAGTGGCGCGCCCTC 661
Qy 762 CTTCGTGGGGGACCGGCTCCACACACACATCGCGCCCTAGAGAGCGCGCCAT 821
Db 662 CTTCGTGGGGGACCGGCTCCACACACACATCGCGCCCTAGAGAGCGCGCCAT 721
Qy 822 CTGGAGCAAGAGAGGATTAACAGAAAGGACACCTCTC 861
Db 722 CTGGAGCAAGAGAGGATTAACAGAAAGGACACCTCTC 761
RESULT 10
ID AAA75165 standard; cDNA; 969 BP.
AC AAA75165;
XX
DT 15-JAN-2001 (first entry)
XX
DE cDNA clone encoding a human TANGO 261 polypeptide.
XX
KW TANGO 266; TANGO 261; TANGO 262; TANGO 267;
KW cellular proliferation; cellular differentiation; cellular adhesion;
KW von Willebrand factor-associated disorder; cell trafficking; cancer;
KW hematopoietic associated disease; atelectasis; pulmonary congestion;
KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;
KW intestinal disorder; spleen associated disease; renal disorder;
KW cardiovascular disorder; ischemic heart disease; hydrocephalus;
KW brain herniation; iatrogenic disease; inflammation; meningitis;
KW Alzheimer's disease; cerebral toxoplasmosis; Parkinson's disease;
KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.

OS Homo sapiens.
XX Key Location/Qualifiers
FH CDS 6..764
FT /**tag= a
FT /product= "TANGO 261"
XX WC200052022-AL.
XX 08-SEP-2000.
XX 01-MAR-2000; 2000WO-US005226.
XX 01-MAR-1999; 99US-0122458P.
XX (MILL-) MILLENNIUM PHARM INC.
XX Barnes TM, Holtzman DA, Sharp JD, Fraser CC;
XX WPI; 2000-579269/54.
XX P-PSDB; AAB18463.
XX Novel human and murine secreted proteins designated TANGO 216, 261, 262,
XX 266 and 267 useful as modulating agents of cellular processes, e.g. for
XX treating cancer.
XX Disclosure; Page; 175pp; English.
XX AAA75163-65 encode human TANGO 261 proteins. The specification also
XX describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO
XX polypeptides can be used to modulate cellular proliferation, modulate
XX cellular differentiation and/or modulate cellular adhesion. The proteins
XX can be used to treat any von Willebrand factor-associated disorder,
XX regulate extracellular matrix structuring, cellular adhesion, and cell
XX trafficking and/or migration, modulate cellular interactions, modulate the
XX cell adhesion in proliferative disorders, such as cancer, modulate the
XX proliferation, differentiation, and/or function of cells that appear in
XX the bone marrow, and leukocytes, treat bone marrow, blood and
XX hematopoietic associated diseases and disorders, atelectasis, pulmonary
XX congestion or oedema, emphysema, chronic bronchitis, bronchial asthma and
XX bronchiectasis, intestinal disorders, spleen associated diseases, such as
XX moderate renal disorders, treat cardiovascular disorders such as ischemic
XX heart disease, modulate the proliferation, differentiation, and/or
XX function of bone and cartilage cells and to treat bone and/or cartilage
XX associated diseases or disorder. They may also be used to treat disorders
XX associated with the ovaries, and cerebral oedema, hydrocephalus, brain
XX herniations, iatrogenic disease, inflammations, bacterial and viral
XX meningitis, Alzheimer's Disease, cerebral toxoplasmosis, Parkinson's
XX disease, multiple sclerosis, brain cancers, hydrocephalus and
XX encephalitis, and treat hepatic disorders. note: the present sequence
XX does not appear in the specification; it was created using information
XX provided
XX
SQ Sequence 969 BP; 210 A; 317 C; 294 G; 148 T; 0 U; 0 Other;
Query Match 87.5%; Score 753.6; DB 3; Length 969;
Best Local Similarity 99.5%; Pred. No. 1.9e-145;
Matches 756; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 102 GGAGATCGGAGCTGGAGTTTGTTCAGATCATCATCGTGTGTGATGATGATGAT 161
Db 2 GGAGATCGGAGCTGGAGTTTGTTCAGATCATCATCGTGTGTGATGATGATGATGAT 61
Qy 162 GTGGTGGTGTGATCAGTGGCTGTGAGCCACTAAGAGCTGTCTGACGGTCTTTCATCAG 221
Db 62 GTGGTGGTGTGATCAGTGGCTGTGAGCCACTAAGAGCTGTCTGACGGTCTTTCATCAG 121
Qy 222 CCGGCACAGCCAGGGCGGAGAGAGTGCCTGTCTCAGAGGATGCTGTGGCC 281
Db 122 CCGGCACAGCCAGGGCGGAGAGAGTGCCTGTCTCAGAGGATGCTGTGGCC 181
Qy 282 CTCGGAGAGCACAGTGTTCAGGCAACGGAATCCAGAGCCGACGCTCTACGCCCGCTCG 341

FT /*tag= a
XX /product= "Neuron associated protein"
PN
XX
XX WO200034477-A2.
PD
XX 15-JUN-2000.
XX
XX 10-DEC-1999; 99WO-US030408.
XX
XX 11-DEC-1998; 98US-00210083.
PR 09-FEB-1999; 99US-0119365P.
PR 16-MAR-1999; 99US-0124687P.
XX
XX (INCY-) INCYTE PHARM INC.
PA
XX
XX Tang YT, Yue H, Baughn MR, Hillman JL, Lal P, Au-Young J;
PI Yang J, Lu DAM, Azimzai Y;
XX
XX WPI: 2000-423423/36.
DR P-PSDB; AAB01398.
DR
XX

New human neuron-associated proteins and polynucleotides encoding them, useful for diagnosis, treatment and prevention of cell proliferative disorders including cancer, neuronal and neurological disorders.

Claim 9; Page 136; 145pp; English.

Human neuron-associated proteins (NEUPAP) can be used for treating or preventing a disorder associated with decreased expression or activity of NEUPAP. Antagonists of NEUPAP are useful for treating or preventing disorder associated with increased expression or activity of NEUPAP or their fragments or derivatives are useful for treating neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia and Parkinson's disease. NEUPAPs are also useful for treating other demyelinating diseases, bacterial and viral meningitis, prion diseases including kuru, Creutzfeldt-Jakob disease, nutritional and metabolic diseases of the nervous system, neurofibromatosis, other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, inherited metabolic, endocrine, and toxic myopathies, mental disorders including mood, anxiety and schizophrenic disorders, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal haemoglobinuria, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, and an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal parasitic, protozoal, and helminthic infections, and trauma. This sequence was given the Incyte ID no. 1871288CB1

XX Sequence 1061 BP; 225 A; 342 C; 326 G; 169 T; 0 U; 0 Other;

Query Match 87.4%; Score 752.6; DB 3; Length 1061;
Best Local Similarity 98.2%; Pred. NO. 3.1e-145;
Matches 761; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 87 TTGTTCAGAGATGAGATACCGAGCTGGAGTTGTTCAGATCATCATCATCGTGT 146
DB 82 TCTCCCGGAAACACGCAATGCGGAGCTGGAGTTGTTCAGATCATCATCATCGTGT 141
QY 147 GGTGATGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 206
DB 142 GGTGATGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 201
QY 207 ACGGTCTTTCATAGCCCGACACCCAGCGCGGAGAGAGATGCCCTGCTCTCAGA 266
DB 202 ACGGTCTTTCATAGCCCGACACCCAGCGCGGAGAGAGATGCCCTGCTCTCAGA 261

QY 267 AGGATGCTGTGCGCCCTCGGAGAGACAGTGTTCAGGCAACGGAATCCAGAGCCGCGAGT 326
DB 262 AGGATGCTGTGCGCCCTCGGAGAGACAGTGTTCAGGCAACGGAATCCAGAGCCGCGAGT 321
QY 327 CTACGCGCGCGCTCGGCGCCACCGAGCGCTGCGCGCTTTCGCGCCACGCGGAGCG 386
DB 322 CTACGCGCGCGCTCGGCGCCACCGAGCGCTGCGCGCTTTCGCGCCACGCGGAGCG 381
QY 387 CTTTCCACCGCTTTCAGCGCCACCTATCCGTAACCTGAGACAGAGATGAGCTGCCACCCAC 446
DB 382 CTTTCCACCGCTTTCAGCGCCACCTATCCGTAACCTGAGACAGAGATGAGCTGCCACCCAC 441
QY 447 CATCTCGCTGTACAGCGGAGAGCGCCACCTACAGGGCGCCCTGTCACCTCCAGCT 506
DB 442 CATCTCGCTGTACAGCGGAGAGCGCCACCTACAGGGCGCCCTGTCACCTCCAGCT 501
QY 507 TCGGGAACCCCGAGCAGCTGGAACCTGAACCGGAGTGGTGGCGCACCCCAACAG 566
DB 502 TCGGGAACCCCGAGCAGCTGGAACCTGAACCGGAGTGGTGGCGCACCCCAACAG 561
QY 567 AACCATCTTCGACAGTACCTGATGATAGTCCAGGCTGGGGCGCCCTGCCCCCAG 626
DB 562 AACCATCTTCGACAGTACCTGATGATAGTGGCGGCTGGGGCGCCCTGCCCCCAG 621
QY 627 CAGTAACTCGGGCATCAGCGCCACCTGCTACCGCAGCGCGCGCGCATGAGGGGCGCC 686
DB 622 CAGTAACTCGGGCATCAGCGCCACCTGCTACCGCAGCGCGCGCGCATGAGGGGCGCC 681
QY 687 GCCCAGCTACAGGAGTATCGGCACCTACCGGGGTCTCTTCCAGACACGAGAG 746
DB 682 GCCCAGCTACAGGAGTATCGGCACCTACCGGGGTCTCTTCCAGACACGAGAG 741
QY 747 CAGTGGCGCGCCCTCTCTGCTGGAGGGGACCCGGGCTCCACACACACATGCGGCCCT 806
DB 742 CAGTGGCGCGCCCTCTCTGCTGGAGGGGACCCGGGCTCCACACACACATGCGGCCCT 801
QY 807 AGAGAGCGCAGCATCTGAGCAGCAAGAGAGAGATGAGTAAAGAGAGACACCTCTC 861
DB 802 AGAGAGCGCAGCATCTGAGCAGCAAGAGAGAGATGAGTAAAGAGAGACACCTCTC 856

RESULT 14
ADC37324
ID ADC37324 standard; DNA; 1085 BP.
XX
XX ADC37324;
XX
XX 18-DEC-2003 (first entry)
XX
XX Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 157.
XX
XX Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
XX cancer; infectious disease; bone disease; AIDS;
XX neurodegenerative disease; ischaemic disorder; Antiinflammatory;
XX Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
XX Neuroprotective; Nootropic; Cardiant; Gene therapy; human; gene; ds.
XX
XX Homo sapiens.
OS
XX WO2003048202-A2.
XX
XX 12-JUN-2003.
XX
XX 03-DEC-2002; 2002WO-JP012644.
XX
XX 03-DEC-2001; 2001JP-00368692.
XX
XX 05-DEC-2001; 2001US-0335829P.
XX
XX 04-OCT-2002; 2002JP-00291302.
XX
XX 04-OCT-2002; 2002US-0415769P.
XX
XX (ASAH) ASAH KASEI KK.
XX

PI Matsuda A, Muramatsu S;
XX WPI: 2003-505282/47.
DR P-PSDB; ADC37325.
XX
PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
PT useful for treating inflammation, autoimmune diseases, cancers,
PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
PT ischemic disorders.
XX
PS Claim 4; SEQ ID NO 157; 938pp; English.
XX
CC The present invention relates to novel proteins and their coding
CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
CC kappaB). The proteins and their coding sequences are useful for treating
CC a disease associated with NF-kappaB activation, such as inflammation,
CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
CC neurodegenerative diseases, or ischaemic disorders.
XX
SQ Sequence 1085 BP; 223 A; 352 C; 334 G; 176 T; 0 U; 0 Other;
Query Match 87.4%; Score 752.6; DB 9; Length 1085;
Best Local Similarity 98.2%; Freq. No. 3.1e-145;
Matches 761; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 87 TTGTGTTCCAGAGCATGGAGATCAGCGAGCTGGAGTTGTTTCAGATCATCATCATCGTGT 146
Db 109 TCTCTCGGAACACGAGCAATGCGGAGCTGGAGTTGTTTCAGATCATCATCATCGTGT 168
QY 147 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 206
Db 169 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 228
QY 207 ACGGTCTCTTCATCAGCGCGGCACAGCAGCGCGGCGGAGAGAGAGAGAGAGAGAGAGAG 266
Db 229 ACGGTCTCTTCATCAGCGCGGCACAGCAGCGCGGCGGAGAGAGAGAGAGAGAGAGAGAG 288
QY 267 AGGATGCTGTGCTCTCGGAG 326
Db 289 AGGATGCTGTGCTCTCGGAG 348
QY 327 CTACGCGCGCGCTCGCGCCACAGCAGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 386
Db 349 CTACGCGCGCGCTCGCGCCACAGCAGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 408
QY 387 CTTTCCACCGCTTCAGCGCCACCTATCCGTACCTGACGACGAGATCGACCTGCCACCCAC 446
Db 409 CTTTCCACCGCTTCAGCGCCACCTATCCGTACCTGACGACGAGATCGACCTGCCACCCAC 468
QY 447 CATCTCGCTGACAGCGGGAGAGCGCCACCTACCGGGCGCGCGCGCGCGCGCGCGCGCGCG 506
Db 469 CATCTCGCTGACAGCGGGAGAGCGCCACCTACCGGGCGCGCGCGCGCGCGCGCGCGCGCG 528
QY 507 TCGGGACCCCGAGCAGCTGGAACCTGAACCGGGAGTGGTGGCGCGCACCCCGCCAAACAG 566
Db 529 TCGGGACCCCGAGCAGCTGGAACCTGAACCGGGAGTGGTGGCGCGCACCCCGCCAAACAG 588
QY 567 AACCATCTTCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 626
Db 589 AACCATCTTCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 648
QY 627 CAGTAACTCGGGCATCAGCGCCACCTGCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 686
Db 649 CAGTAACTCGGGCATCAGCGCCACCTGCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 708
QY 687 GCCACCTACAGGAGTGAATGGGCGAATACCGGGGGTCTCTTTCAGACACAGCAGAG 746
Db 709 GCCACCTACAGGAGTGAATGGGCGAATACCGGGGGTCTCTTTCAGACACAGCAGAG 768
QY 747 CAGTGGGCGCGCTCTCTTGTGGAGGGGACCGGGCTCCACACACACACATCGCGCCCT 806
Db 769 CAGTGGGCGCGCTCTCTTGTGGAGGGGACCGGGCTCCACACACACACATCGCGCCCT 828

QY 807 AGAGAGCGCGCCATCTGGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 861
Db 829 AGAGAGCGCGCCATCTGGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 883
RESULT 15
ABZ36103
ID ABZ36103 standard; cDNA; 1334 BP.
XX
AC ABZ36103;
XX
DT 10-FEB-2003 (first entry)
XX
DE Human secretory polynucleotide SPTM SEQ ID NO 267.
XX
KW Human; SPTM; autoimmune disorder; inflammatory disorder; AIDS; anaemia;
KW asthma; Crohn's disease; neurological disorder; epilepsy; cancer;
KW Huntington's disease; Alzheimer's disease; Creutzfeldt-Jakob disease;
KW multiple sclerosis; Parkinson's disease; cell proliferative disorder;
KW anti-inflammatory; immunosuppressive; neuroprotective; nootropic;
KW neuroleptic; anticonvulsant; cytostatic; antiparkinsonian; anxiolytic;
KW antipsoriatic; antianaemic; anti-HIV; human immunodeficiency virus;
KW secretory polynucleotide; secretory protein; gene; ss.
OS Homo sapiens.
XX
PN WO200283876-A2.
XX
PD 24-OCT-2002.
XX
PF 27-MAR-2002; 2002WO-US009921.
XX
PR 29-MAR-2001; 2001US-0280067P.
PR 29-MAR-2001; 2001US-0280068P.
PR 16-MAY-2001; 2001US-0291280P.
PR 17-MAY-2001; 2001US-0291829P.
PR 17-MAY-2001; 2001US-0291849P.
PR 19-JUN-2001; 2001US-0299428P.
PR 20-JUN-2001; 2001US-0299776P.
PR 20-JUN-2001; 2001US-0300010P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshay SR;
PI Daughtery SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstein EH;
PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
PI Flores V, Marwaha R, Lo A, Lan RV, Urashka ME;
XX
DR WPI: 2003-075543/07.
DR P-PSDB; ABP75660.
XX
PT New human secretory proteins and polynucleotides, useful for diagnosing,
PT treating or preventing autoimmune/inflammatory disorders (e.g. AIDS),
PT neurological disorders (e.g. Alzheimer's), or cell proliferations or
PT cancers.
XX
PS Claim 1; SEQ ID NO 267; 458pp + Sequence Listing; English.
XX
CC The invention relates to a secretory polynucleotide (designated sptm)
CC comprising any of 567 polynucleotide sequences (ABZ35837-ABZ36403), a
CC naturally occurring polynucleotide sequence at least 90 % identical to
CC the polynucleotide sequence, a polynucleotide complementary to them or an
CC RNA equivalent of them. The polypeptide or polynucleotide are useful for
CC treating, preventing or diagnosing a disease or condition associated with
CC the expression of functional SPTM. These are particularly useful for
CC the diagnosis, treating or preventing autoimmune/inflammatory disorders
CC (e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's
CC disease), neurological disorders (e.g. epilepsy, Huntington's disease,
CC dementia, stroke, Alzheimer's disease, Creutzfeldt-Jakob disease,
CC multiple sclerosis, cerebral palsy, Parkinson's disease, anxiety,
CC schizophrenia or amnesia), or cell proliferative disorders (e.g.
CC psoriasis, polycythemia vera, or cancers including adenocarcinoma,

CC leukaemia, lymphoma, melanoma, myeloma, sarcoma or cancers of the brain,
CC breast, cervix or prostate). Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in electronic
XX format directly from WIPO at ftp.wipo.int/pub/published_pcr_sequences
SQ Sequence 1334 BP; 299 A; 381 C; 394 G; 260 T; 0 U; 0 Other;

Query Match 87.4%; Score 752.6; DB 7; Length 1334;
Best Local Similarity 98.2%; Pred. No. 3.2e-145;
Matches 761; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 87 TTGTTCAGAGCATGGAGATCAGCGAGCTGGAGTTGTTTCAGATCATCATCATCGTGGT 146
Db 82 TCTCTGCGAACACAGGAATGGCGAGCTGGAGTTGTTTCAGATCATCATCATCGTGGT 141
QY 147 GGTGATGATGGTATGGTGGTGGTATCATCGTGCCTGCTGAGCCACTACAAAGCTGTCTGC 206
Db 142 GGTGATGATGGTATGGTGGTGGTATCATCGTGCCTGCTGAGCCACTACAAAGCTGTCTGC 201
QY 207 AGGTCCTTTCATCAGCCGCGCAGCCAGCGGGGAGGAGAGAGATGCCCTGTCTCTAGA 266
Db 202 AGGTCCTTTCATCAGCCGCGCAGCCAGCGGGGAGGAGAGAGATGCCCTGTCTCTAGA 261
QY 267 AGGATGCTGTGGCCCTCGAGAGCACAGTGTACGGCAACGGAAATCCAGAGCCGCGAGGT 326
Db 262 AGGATGCTGTGGCCCTCGAGAGCACAGTGTACGGCAACGGAAATCCAGAGCCGCGAGGT 321
QY 327 CTACGCCCGCTCGGCCACCGACCGCTGCGTGCCTGCGCCAGCGGGAGCG 386
Db 322 CTACGCCCGCTCGGCCACCGACCGCTGCGTGCCTGCGCCAGCGGGAGCG 381
QY 387 TTTCACCGCTTTCAGCCGCTATCCCTTACCTGACGACGAGATCGACCTGCCACCCAC 446
Db 382 TTTCACCGCTTTCAGCCGCTATCCCTTACCTGACGACGAGATCGACCTGCCACCCAC 441
QY 447 CATCTCGTGTACAGGGGAGAGCCGCCCTACAGGGCCCTGACCCCTCCAGCT 506
Db 442 CATCTCGTGTACAGGGGAGAGCCGCCCTACAGGGCCCTGACCCCTCCAGCT 501
QY 507 TCGGACCCCGAGCAGCTGGAATGAACCGGAGTCTGGTGCAGCCACCCCAACAG 566
Db 502 TCGGACCCCGAGCAGCTGGAATGAACCGGAGTCTGGTGCAGCCACCCCAACAG 561
QY 567 AACCATCTTCGACAGTACCTGATGGATAGTCCAGGCTGGGCGGCCCTGCCCCCAG 626
Db 562 AACCATCTTCGACAGTACCTGATGGATAGTCCAGGCTGGGCGGCCCTGCCCCCAG 621
QY 627 CAGTAACTCGGCGCATCAGCGCCACCTGCTACGGCAGCGCGCGCATGGAGGGCGGCC 686
Db 622 CAGTAACTCGGCGCATCAGCGCCACCTGCTACGGCAGCGCGCGCATGGAGGGCGGCC 681
QY 687 GCCCACTACAGCGAGGTATCGGCCATACCCGGGGTCTCTCTCCAGCACCCAGCAG 746
Db 682 GCCCACTACAGCGAGGTATCGGCCATACCCGGGGTCTCTCTCCAGCACCCAGCAG 741
QY 747 CAGTGGGCGCCCTCTCTGAGGGGACCGCGCTCCACACACACATCGGCCCT 806
Db 742 CAGTGGGCGCCCTCTCTGAGGGGACCGCGCTCCACACACACATCGGCCCT 801
QY 807 AGAGCGGAGCCATCTGGAGCAAGAGAGAGATAAACAGAAAGGACCCCTCTC 861
Db 802 AGAGCGGAGCCATCTGGAGCAAGAGAGAGATAAACAGAAAGGACCCCTCTC 856

Search completed: May 12, 2004, 00:56:47

Job time : 395.646 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2004, 00:10:42 ; Search time 2585.37 Seconds
(without alignments)
9944.945 Million cell updates/sec

Title: US-09-934-249-3

Perfect score: 861

Sequence: 1 atgcaccgcttgatgggggt.....aacagaaaggacaccccttc 861

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:**
1: em_estba:**
2: em_esthum:**
3: em_estlin:**
4: em_estnu:**
5: em_estov:**
6: em_estpl:**
7: em_estro:**
8: em_htc:**
9: gb_est1:**
10: gb_est2:**
11: gb_htc:**
12: gb_est3:**
13: gb_est4:**
14: gb_est5:**
15: em_estfun:**
16: em_estcom:**
17: em_gss_hum:**
18: em_gss_hiv:**
19: em_gss_pin:**
20: em_gss_vrt:**
21: em_gss_fun:**
22: em_gss_mam:**
23: em_gss_mus:**
24: em_gss_pro:**
25: em_gss_rod:**
26: em_gss_phg:**
27: em_gss_vrl:**
28: gb_gss1:**
29: gb_gss2:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	811.8	94.3	1009	9	AL578575
c 2	807.6	93.8	1201	9	AL517150
c 3	743.6	86.4	967	13	BQ641849
4	700	81.3	1046	12	BM922276

5	687.8	79.9	945	13	BU539219
6	687	79.8	951	9	AL558881
c 7	667.4	77.5	916	13	BQ954555
8	667.4	77.5	1007	9	AL558882
c 9	667	77.5	874	13	BX362396
10	662.8	77.0	850	13	BU602918
11	648	77.3	780	29	AY419334
12	614.6	71.4	1079	11	BC023092
13	607.4	70.5	609	13	BQ636742
14	603.6	70.1	1207	11	AK008976
15	578.4	67.2	890	13	BQ690750
16	570.4	66.2	973	13	BU169156
17	568.8	66.1	572	13	BX641317
c 18	568.2	66.0	782	12	BQ015170
19	550	63.9	551	12	BM141979
20	529.8	61.5	894	12	BI851941
c 21	519.4	60.3	729	13	BQ575741
c 22	516.8	60.0	730	12	BM677602
c 23	515.2	59.8	728	13	BU683523
24	504	58.5	780	29	AY419335
25	501.6	58.3	588	12	BM483503
26	487.8	56.7	1201	9	AL543170
27	487.6	56.6	646	29	CG784226
28	475.2	55.2	624	29	AY419336
29	468	54.4	857	12	BG323347
c 30	465.8	54.1	744	13	BU414421
31	461	53.5	763	12	BI646175
c 32	445.6	51.8	693	9	AI761441
33	445.4	51.7	655	13	BQ691705
34	445.4	51.7	1280	13	BQ691500
35	443.6	51.5	651	14	CB554226
36	440	51.1	899	13	BU136912
37	432.8	50.3	964	13	BQ859860
38	430	49.9	841	12	BI156703
39	429.6	49.9	646	13	BU859841
40	428.8	49.8	974	10	BB624904
41	428.2	49.7	655	12	BI853324
42	416.2	48.3	587	12	BI083462
43	411.6	47.8	1127	13	BU174654
44	411	47.7	582	10	BE553323
c 45	407.4	47.3	629	13	BU730650

ALIGNMENTS

RESULT 1
AL578575/c
LOCUS
DEFINITION AL578575 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
cdna clone CS0DK001VC24 3-PRIME, mRNA sequence.
ACCESSION AL578575
VERSION AL578575.2 GI:31316780
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1009)
Li W.B., Gruber C., Jessee J., and Polayes D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 16, 2001 this sequence version replaced gi:12942781.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9945.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DK001BB12NP1&cluster=9945.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :

BU539219 AGENCOURT
AL558881 AL558881
BQ954555 AGENCOURT
AL558882 AL558882
BX362396 BX362396
BU602918 AGENCOURT
AY419334 Homo sapi
BQ636742 hdl3h06.Y
AK008976 Mus muscu
BQ690750 AGENCOURT
BU169156 AGENCOURT
BX641317 DXF2P686K
BQ015170 UI-H-BD1-
BM141979 if25a11.Y
BI851941 60379004
BQ575741 UI-H-E21-
BM677602 UI-E-E01-
BU683523 UI-CF-EC1
AY419335 Pan trogl
BM483503 536869 MA
AL543170 AL543170
CG784226 FHRCR-GT-
AY419336 Mus muscu
BG323347 602421734
BU414421 603670223
BI646175 603276395
AI761441 wg65f07.x
BQ691705 AGENCOURT
BQ691500 AGENCOURT
CB554226 WMSP0052
BU136912 AGENCOURT
BU859860 AGENCOURT
BI156703 602922119
BU859841 AGENCOURT
BB624904 BB624904
BI853324 60379903
BI083462 602875788
BU174654 AGENCOURT
BE553323 ur44h02.Y
BU730650 UI-E-C11-


```
QY 256 CTCTCTCAGAGGATGCTGTGGCCCTCGGAGACACAGTGTACGGCAACGGAATCCCA 315
Db 770 CTGTCTCTCAGAGGATGCTGTGGCCCTCGGAGACACAGTGTACGGCAACGGAATCCCA 711
QY 316 GAGCCGAGGTCTACGCCCGGCTCGGCCACCGACCGCGTGGCGCCCTTCGCC 375
Db 710 GAGCCGCA-GTCTACGCCCGGCTCGGCCACCGACCGCGTGGCGCCCTTCGCC 652
QY 376 GAGCGGAGCGCTTCCACCGCTTCCAGCCACCTATCCGTACTGTCAGCACAGATCCGAC 435
Db 651 GAGCGGAGCGCTTCCACCGCTTCCAGCCACCTATCCGTACTGTCAGCACAGATCCGAC 592
QY 436 CTGCCACCAACCATCTCGTGTCTCAGACGGGAGGAGCCGCCACCTTACAGGGCCCTTCG 495
Db 591 CTGCCACCAACCATCTCGTGTCTCAGACGGGAGGAGCCGCCACCTTACAGGGCCCTTCG 532
QY 496 ACCTCTCAGTTCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGAGTCTGTCGCCGA 555
Db 531 ACCTCTCAGTTCGGGACCCCGAGCAGCAGCTGGAACCGGAGTCTGTCGCCGA 472
QY 556 CCCCCAAACAGAACCATCTTCGACAGTGCCTGATGATAGTGCCAGGTGGCGGCCGCC 615
Db 471 CCCCCAAACAGAACCATCTTCGACAGTGCCTGATGATAGTGCCAGGTGGCGGCCGCC 412
QY 616 TGCCCCCAGCAGTAACTTCGGGCATCAGGCCACCTGTCAGCGGAGCGGGGCGCATG 675
Db 411 TGCCCCCAGCAGTAACTTCGGGCATCAGGCCACCTGTCAGCGGAGCGGGGCGCATG 352
QY 676 GAGGGGCGCGGCCACCTACAGCAGGTCTACCGGCCATCGGCCACCTACCGGGGTCCTCTTCCAG 735
Db 351 GAGGGGCGCGGCCACCTACAGCAGGTCTACCGGCCATCGGCCACCTACCGGGGTCCTCTTCCAG 292
QY 736 CACAGCAGACAGTGGCGCGCCCTCTGCTGAGAGGAGACCGCGTCCACACACACAC 795
Db 291 CACAGCAGACAGTGGCGCGCCCTCTGCTGAGAGGAGACCGCGTCCACACACACAC 232
QY 796 ATCGCGCCCTAGAGAGCGAGCATCTGAGCAAGAGAGATAAACAGAAAGGACAC 855
Db 231 GTGGCGCCCTAGAGAGCGAGCATCTGAGCAAGAGAGATAAACAGAAAGGACAC 172
QY 856 CTTCTC 861
Db 171 CTTCTC 166

RESULT 3
B0641849
LOCUS
DEFINITION B0641849 967 bp mRNA linear EST 15-JUL-2002
5', mRNA sequence.
ACCESSION B0641849
VERSION B0641849.1 GI:21766021
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
1 (bases 1 to 967)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2493 row: g column: 18
High quality sequence stop: 571.
Location/Qualifiers
```

```
source 1..967
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6292265"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC 43"
/notes="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. |"

ORIGIN
Query Match 86.4%; Score 743.6; DB 13; Length 967;
Best Local Similarity 99.5%; Pred. No. 6.8e-127;
Matches 746; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 112 GAGCTGAGGTTTGTTCAGATCATCATCGTGGTGTGATGATGGTGTGTTGGTGTG 171
Db 1 GAGCTGAGGTTTGTTCAGATCATCATCGTGGTGTGATGATGGTGTGTTGGTGTG 60
QY 172 ATCAGTGTCTGTGAGCCACTCAAGCTGTCTGCAAGTCTTTCATCAGCCGCGCACAGC 231
Db 61 ATCAGTGTCTGTGAGCCACTCAAGCTGTCTGCAAGTCTTTCATCAGCCGCGCACAGC 120
QY 232 CAGGGCGGAGGAGAGAGATGCCCTGTCTCAAGAGGATGCCCTGTGGCCCTCGGAGAGC 291
Db 121 CAGGGCGGAGGAGAGAGATGCCCTGTCTCAAGAGGATGCCCTGTGGCCCTCGGAGAGC 180
QY 292 ACAGTGTACGCAACGGAATCCAGAGCCGCGAGTCTTACGCCCGCCCTTGGCCCGCACCGAC 351
Db 181 ACAGTGTACGCAACGGAATCCAGAGCCGCGAGTCTTACGCCCGCCCTTGGCCCGCACCGAC 240
QY 352 CGCTTGGCGTGGCGCCCTTCCGCCAGCGGAGCGTTCACCGCTTCCAGCCCGCACCTAT 411
Db 241 CGCTTGGCGTGGCGCCCTTCCGCCAGCGGAGCGTTCACCGCTTCCAGCCCGCACCTAT 300
QY 412 CCGTACCTTGACGACGAGATCGACCTGCAACCCACCATCTCGTGTGTCAGAGGGGAGAG 471
Db 301 CCGTACCTTGACGACGAGATCGACCTGCAACCCACCATCTCGTGTGTCAGAGGGGAGAG 360
QY 472 CCCCACCTTACGAGGGCCCTTGCACCTTCAGCTTGGGACCCCGGAGCAGCGCTGAA 531
Db 361 CCCCACCTTACGAGGGCCCTTGCACCTTCAGCTTGGGACCCCGGAGCAGCGCTGAA 420
QY 532 CTGAACCGGAGTGGTGGCGGCGCATCGAGGGCGCGCCCAACAGAACCATCTTTCGACAGTGCCTGATG 591
Db 421 CTGAACCGGAGTGGTGGCGGCGCATCGAGGGCGCGCCCAACAGAACCATCTTTCGACAGTGCCTGATG 480
QY 592 GATAGTGCAGGCTGGGCGGCCCTTGCCTTCCAGACCGAGTAACTCGGGATCAGCGCCACG 651
Db 481 GATAGTGCAGGCTGGGCGGCCCTTGCCTTCCAGACCGAGTAACTCGGGATCAGCGCCACG 540
QY 652 TGCTACGCGACGCGGCGCGCATCGAGGGCGCGCCCGCCACCTACAGCGAGGTTCATCGGC 711
Db 541 TGCTACGCGACGCGGCGCGCATCGAGGGCGCGCCCGCCACCTACAGCGAGGTTCATCGGC 600
QY 712 CACTACCGGGGTCTCTCTTCCAGACCGAGCAGTGGGCGCCCTCTCTTCTGCTGAG 771
Db 601 CACTACCGGGGTCTCTCTTCCAGACCGAGCAGTGGGCGCCCTCTCTTCTGCTGAG 660
QY 772 GGGACCGGGTCCACACACACATCGCGCCCTTAGAGAGCGGAGCCATCTGAGGACAA 831
Db 661 GGGACCGGGTCCACACACACATCGCGCCCTTAGAGAGCGGAGCCATCTGAGGACAA 720
QY 832 GAGAAGGATAAACAGAAAGGACACCCCTCTC 861
Db 721 GAGAAGGATAAACAGAAAGGACACCCCTCTC 750
```

[illegible]

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9945.r For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DJ015DC06Q1&cluster=9945.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DJ015DC06Q1.
Location/Qualifiers
1. .1007
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ015VF12"
/call_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
source
Query Match 77.5%; Score 667.4; DB 9; Length 1007;
Best Local Similarity 99.0%; Pred. No. 7.6e-113;
Matches 690; Conservative 2; Mismatches 3; Indels 2; Gaps 2;
Qy 165 GGTGGTATCAGTGCCTGCTGAGCCACTACAGTGTCTGACGCTCTCTCATCAGCG 224
Db 62 GGATGTATCAGTGCCTGCTGAGCCACTACAGTGTCTGACGCTCTCTCATCAGCG 121
Qy 225 GCACAGCAGCGGCGGAGGAGAGATGCCCTGTCTCTCAGAGGATGCCCTGTGGCCCTC 284
Db 122 GCACAGCAGCGGCGGAGGAGAGATGCCCTGTCTCTCAGAGGATGCCCTGTGGCCCTC 181
Qy 285 GGAGAGACAGTGTACAGCAACGAAATCCAGAGCGCAGGTCTACGCCGCTCGGCC 344
Db 182 GGAGAGACAGTGTACAGCAACGAAATCCAGAGCGCAGCA-GTCTACGCCGCTCGGCC 240
Qy 345 CACCGACCGCTGCGGCGCTTGGCCAGCGGAGCGGTCTCCACCGCTTCCAGCC 404
Db 241 CACCGACCGCTGCGGCGCTTGGCCAGCGGAGCGGTCTCCACCGCTTCCAGCC 300
Qy 405 CACTATCGTACCTGAGCAGAGATGACCTGCCACCAACCATCTGCTGTACAGCG 464
Db 301 CACTATCGTACCTGAGCAGAGATGACCTGCCACCAACCATCTGCTGTACAGCG 360
Qy 465 GGAGAGCGCCACCTTACAGAGCGCCCTGACCTCCAGCTTGGAGACCCCGAGCAGCA 524
Db 361 GGAGAGCGCCACCTTACAGAGCGCCCTGACCTCCAGCTTGGAGACCCCGAGCAGCA 420
Qy 525 GCTGGAATGAAACCGGAGTGGTGGCGGCAACCCCAACAGAACCATCTTCAGCAGTGA 584
Db 421 GCTGGAATGAAACCGGAGTGGTGGCGGCAACCCCAACAGAACCATCTTCAGCAGTGA 480
Qy 585 CCTGATGATAGTGCAGGCTGGGCGGCGCTGCGCCCGCAGAGTAACTCGGGCATCAG 644
Db 481 CCTGATGATAGTGCAGGCTGGGCGGCGCTGCGCCCGCAGAGTAACTCGGGCATCAG 540
Qy 645 CGCCACGTGTACCGGAGCGGCGGCGATGGAGGGGCGCGCCACCTACAGCGAGGT 704
Db 541 CGCCACGTGTACCGGAGCGGCGGCGATGGAGGGGCGCGCCACCTACAGCGAGGT 600
Qy 705 CATGGCCCACTACCGGGGTCTCTTCCAGCACCAGCAGAGTGGGCGGCGCTCTT 764
Db 601 CATGGCCCACTACCGGGGTCTCTTCCAGCACCAGCAGAGTGGGCGGCGCTCTT 660
Qy 765 GCTGAGGGGACCCGGCTTCCACACACACATCGCGCCCTTAGAGCGCGAGCCATCTG 824
Db 661 GCTGAGGGGACCCGGCTTCCACACACACATCGCGCCCTTAGAGCGCGCA-SCATCTG 719

Qy 825 GAGCAAGAGAGGATTAACAGAGAGGACACCCCTCTC 861
Db 720 GAGCAAGAGAGGATTAACAGAGAGGACACCCCTCTC 756

RESULT 9
BX362396/c
LOCUS BX362396 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
DEFINITION Homo sapiens cDNA clone CS0DJ014YN15 3-PRIME, mRNA sequence.
ACCESSION BX362396
VERSION BX362396
KEYWORDS BX362396.1 GI:30378625
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 874)
AUTHORS Li.W.B., Gruber.C., Jessee.J. and Polayes.D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9945.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DJ014CG08NP1&cluster=9945.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DJ014CG08NP1.
Location/Qualifiers
1. .874
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ014YN15"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 77.5%; Score 667; DB 13; Length 874;
Best Local Similarity 95.9%; Pred. No. 8.7e-113;
Matches 693; Conservative 16; Mismatches 11; Indels 2; Gaps 2;
Qy 151 ATGATGGTATGGTGGTGGTATC-ACGTGGCTGTGAGCCACTACAGCTGTCTGCAGC 209
Db 874 ATGATGGTATGGTGGTGGTATC-ACGTGGCTGTGAGCCACTACAGCTGTCTGCAGC 815
Qy 210 GTCTCTTATCAGCCGCGCAGCCAGCGGCGGAGAGAGATGCCCTGTCTCAGAGG 269
Db 814 GTCTCTTATCAGCCGCGCAGCCAGCGGCGGAGAGAGATGCCCTGTCTCAGAGG 755
Qy 270 ATGCTGTGGCCCTCGGAGAGCAGATGTTCAGGCAACCGAAATCCAGAGCCCGAGTCTA 329
Db 754 ATGCTGTGGCCCTCGGAGAGCAGATGTTCAGGCAACCGAAATCCAGAGCCCGCA-GTYTA 696
Qy 330 CGCCCGGCTCGGCGCAGCCGCGCTGGCGGCGCTTCCGCCCGGCGGCGGCTT 389
Db 695 CGCCCGGCTCGGCGCAGCCGCGCTGGCGGCGCTTCCGCCCGGCGGCGGCTT 636
Qy 390 CCACCGCTTCCAGCGCCACTTACCTGCTGAGCAGCAGAGATCGACCTGCCACCCACCAT 449
Db 635 CCACCGCTTCCAGCGCCACTTACCTGCTGAGCAGCAGAGATCGACCTGCCACCCACCAT 576

450	Qy	CTCGTGTTCACACGGGAGGAGCCCCACACCTTACGAGGGCCCTGACCCCTCGAGCTTCG	509
575	Db	CTCGTGTTCACACGGGAGGAGCCCCACACCTTACGAGGGCCCTGACCATCAGTCTTCG	516
510	Qy	GGACCCCGACGACGAGCTTGGAACTGAACCGGGAGTCTGTCGCGCACCCCCAAAACAGAAC	569
515	Db	GGACCCCGAGGACGAGCTTGGAACTGAACCGGGAGTCTGTCGCGCACCCCCAAAACAGAAC	456
570	Qy	CATCTTCGACAGTGAACCTGATGGTACTGCGCAGGCTGGCGGCGCCCTGTCCCGCCCGACAG	629
455	Db	CATCTTCGACAGTGAACCTGATGGTACTGCGCAGGCTGGCGGCGCCCTGTCCCGCCCGACAG	396
630	Qy	TAACTCGGGCATCAGCGGCCACGTCGTACGCGAGCGGCGGCGCGATGGAGGGCGCGCCGC	689
395	Db	TAACTCGGGCATCAGCGGCCACGTCGTACGCGAGCGGCGGCGCATGGAGGGCGCGCCGC	336
690	Qy	CACCTACAGCGAGTCACTCGGCCACTACCGGGGTCTCTCTCCACGACACGACAGAGAG	749
335	Db	CACCTACAGCGAGTCACTCGGCCACTACCGGGGTCTCTCTCCACGACACGAGAGAGAG	276
750	Qy	TGGCGCGCCCTCTTGTGTGGAGGGAGCCCGGCTCCACACACATCGCGCCCTTAGA	809
275	Db	TGGCGCGCCCTCTTGTGTGGAGGGAGCCCGGCTCCACCAACACATCGCGCCCTTAGA	216
810	Qy	GAGCGCAGCCATCTTGGAGCAAGAGAAAGGATAAACAGAAAGGACACCTCTC	861
215	Db	GAGCGCAGCCATCTTGGAGCAAGAGAGGATAAACAGAAAGGACACCTCTC	164

RESULT 10	BU602918	850 bp	mRNA	linear	EST 20-SEP-2002
LOCUS	AGNC02918	10016502	NIH_MGC_142	Homo sapiens	cDNA clone
DEFINITION	IMAGS:6497853 5', mRNA_sequence.				
ACCESSION	BU602918				
VERSION	BU602918.1	GI:23254677			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE	1. (bases 1 to 850)				
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/ .				
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)				
	Unpublished (1999)				
	Contact: Robert Strausberg, Ph.D.				
	Email: cgabs@email.nih.gov				
	Tissue procurement: NCI				
	cDNA Library Preparation: Michael Brownstein Laboratory				
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)				
	DNA Sequencing by: Agencourt Bioscience Corporation				
	Clone distribution: MGC clone distribution information can be				
	found through the I.M.A.G.E. Consortium/LLNL at:				
	http://image.llnl.gov				
	Plate: LLCM2679	row: 1	column: 22		
	High quality sequence	stop: 1	column: 22		
					499.

FEATURES
SOURCE

```

1. 850
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6497853"
/tissue_type="mixed (pool of 40 RNAs)"
/lab_host="DH10B (T1-phage-resistant)"
/clone_lib="NIH MGC 142"
/notes="Vector: pDNR-LIB; Site 1: 5'II (ggccattatggcc);
Site 2: 5'II (ggccctctggcc); Double-stranded cDNA was
prepared from a pool of 40 cell line polyA+ RNAs (bladder
- 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -
4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,
kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%,
ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary
gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were

```

ORIGIN

		Query Match Best Local Similarity Matches	77.0%; 97.9%; 703;	Score 662.8; DB 13; Pred.No. S.le-1l2; Conservative	Length 850; Indels 3; Gaps 3;
Qy	147	GGTGATGATGGTGGTGATCGTGATCATCGTGCCTCTGTGAGCCACTACAAGCTGTCTGC	206		
Dd	1	GGTGATGATGGTGGTGATCATCGTGCCTGTGAGCCACTACAAGCTGTCTGC	60		
Qy	207	ACGGTCCTTCATCAGCCGCCACAGCACAGGGGGGAGAGAGATGCCCTGTCCTCAGA	266		
Dd	61	ACGGTCCTTCATCAGCCGGCAGACCAGGGGGGAGAGAGATGCCCTGTCCTCAGA	120		
Qy	267	AGGATGCGCTGTGCCCCTCGAGAGCAACAGTGTGAGCAACGGAAATCCAGAGCCGACGT	326		
Dd	121	AGGATGCGCTGTGCCCCTCGAGAGCAACAGTGTGAGCAACGGAAATCCAGAGCCGACGT	180		
Qy	327	CTAAGCCCCGCTCTGGCCCCACCGAACCGCTGACCGTCCGCGCTTCGCCAGCGGGAGCG	386		
Dd	181	CTAAGCCCCGCTCTGGCCCCACCGAACCGCTGACCGTCCGCGCTTCGCCAGCGGGACCG	240		
Qy	387	CATTCCACCGCTTCACGCCCACTTATCCGTACTCTGCAGCAGCAGATCGACTGCCACCCAC	446		
Dd	241	CATTCCACCGCTTCACGCCCACTTATCOGTACTCTGCAGCAGCAGATCGACTGCCACCCAC	300		
Qy	447	CATCTCGTTGTAGACGGGAGAGCCCCACCTTACACAGGGGCCCTTGACCCCTCCAGCT	506		
Dd	301	CATCTCGTTGTAGACGGGAGAGCCCCACCTTACACAGGGGCCCTTGACCCCTCCAGCT	360		
Qy	507	TCSGGACCCCGACAGCAGCTGGAACTGAACCCGGGAGTCCGTTGGGGCAACCCCAAACAG	565		
Dd	361	TCSGGACCCCGACAGCAGCTGGAACTGAACCCGGGAGTCCGTTGGGGCAACCCCAAACAG	420		
Qy	567	AACATCTTCGACAGTGAACCTGATGATAGTGCAGGCTGGGGGCGCCCTTGCCCCCCCAG	626		
Dd	421	AACATCTTCGACAGTGAACCTGATGATAGTGCAGGCTGGGGGCGCCCTTGCCCCCCCAG	480		
Qy	627	CAGTAACCTCGGGCATCAGCCGCCACTGTCTACGGCAGCGGGCGCGCATGAGAGGGCGCC	686		
Dd	481	CAGTAACCTCGGGCATCAGCCGCCACTGTCTACGGCAGCGGGCGCGCATGAGAGGGCGCC	540		
Qy	687	GCCOACCTACAGCGAGGTCTATCGGGCCAATAACCGGGGTTCCTCTTCAGCACCAGCAGAG	746		
Dd	541	GCCOACCTACAGCGAGGTCTATCGGGCCAATAACCGGGGTTCCTCTTCAGCACCAGCAGAG	600		
Qy	747	CAGTGGGCGC - CCTCTCTTCTCTG - AGGGACCCGGTCCACACACACATCTGGCGCC	804		
Dd	601	CAGTGGGCGCGCCTCTCTCTGAGGGGACCCGGTTCACACACACAAAATCCGGCCC	660		
Qy	805	CTAGAGAGCCGACCA - TCTGGAGCAAGAGAGGATPAACAGAAAGGACACCCCTCTC	861		
Dd	661	CTAAGAGCCGACCAATCTCTGGAGCAGCAGATGATPAACCGAAAGGACCCCTCTC	718		

RESULT 11
AY419334

AV419334
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AV419334
Homo sapiens TM6PAI gene, partial sequence,
genomic survey sequence.
AV419334
AV419334.1
GI:39775291
GSS.
Homo sapiens (human)
Homo sapiens

780 bp
DNA
linear
GSS 17-DEC-2003

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 780)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Smitsky, J.J.,
Adams, M.D. and Cargill, M.
Inferred nonneutral evolution from human-chimp-mouse orthologous
gene tricos
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 780)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Smitsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
These sequences were made by sequencing genomic exons and ordering
them based on alignment.
Location/Qualifiers
1. 780
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..780
/gene="TMEM18"
/locus_tag="HCM6856"
Query Match 75.3%; Score 648; DB 29; Length 780;
Best Local Similarity 86.2%; Pred. No. 2,7e-109;
Matches 648; Conservative 0; Mismatches 104; Indels 0; Gaps 0;
Qy 110 CGAGCTGGAGTTGTTTCAGATCATCATCTGCTGGTGTGATGATGGTGGTGG 169
Db |||||
26 CGAGCTGGAGTTGTTTCAGATCATCATCTGCTGGTGTGATGATGGTGGTGG 85
Qy 170 TGATCAGTGCTGCTCAGCGCAGTACAGAGTGTCTGACCGTCTCTTATCAGCGGCACA 229
Db |||||
86 TGATCAGTGCTGCTCAGCGCAGTACAGAGTGTCTGACCGTCTCTTATCAGCGGCACA 145
Qy 230 GCCAGGCGCGAGAGAGAGATGCCCTGTCTCAGAGAGATGCTGTGCTGCTCGGAGA 289
Db |||||
146 GCCAGGCGCGAGAGAGAGATGCCCTGTCTCAGAGAGATGCTGTGCTGCTCGGAGA 205
Qy 290 GCACAGTGTACGAGCAAGGATCCAGAGCGCGAGGCTACGCGCCCTCGGCCACCG 349
Db |||||
206 GCACAGTGTACGAGCAAGGATCCAGAGCGCGAGGCTACGCGCCCTCGGCCACCG 285
Qy 350 ACCGCTGGCGTGGCGCCCTTCGCCAGCGGAGCGCTTCCACCGCTTCCAGCCACCT 409
Db |||||
266 NNN 325
Qy 410 ATCGTACTCGACAGAGATCGACCTGCGACCCACCATCTCGCTGTACAGCGGAGG 469
Db |||||
326 NNN 385
Qy 470 AGCCCCACCTTACAGGGGCGCCCTGCACCTTCCAGCTTCGCGAGCCCGAGCAGCGTGG 529
Db |||||
386 AGCCCCACCTTACAGGGGCGCCCTGCACCTTCCAGCTTCGCGAGCCCGAGCAGCGTGG 445
Qy 530 AACTGAACCGGAGTGGTGGCGGACCCCAACAGAACATCTTCAGAGTACCTGA 589
Db |||||
446 AACTGAACCGGAGTGGTGGCGGACCCCAACAGAACATCTTCAGAGTACCTGA 505
Qy 590 TGGATAGTGGCAGGCTGGGCGGCGCTGCCCGCCAGCAGTAACTCGGGCATCAGCGCCA 649
Db |||||
506 TGGATAGTGGCAGGCTGGGCGGCGCTGCCCGCCAGCAGTAACTCGGGCATCAGCGCCA 565
Qy 650 CTGCTACGCGAGCGCGGCGCATGAGAGGCGCGCCGCCACCTACAGGAGTCTATCG 709
Db |||||

566 CGTCTACGGCAGCGCGCGCGCATGAGAGGGCGCGCCGCCACCTACAGCGAGTCTATCG 625
Qy 710 GCCACTACCGGGGTCTCTTCCAGCACCAGCAGCAGTGGGCCCTCTCTTGTGG 769
Db |||||
626 GCCACTACCGGGGTCTCTTCCAGCACCAGCAGCAGTGGGCCCTCTCTTGTGG 685
Qy 770 AGGGACCGGGTCTCCACACACACATCGGCCCTAGAGAGCGCAGCATCTGGAGCA 829
Db 686 AGGGACCGGGTCTCCACACACACATCGGCCCTAGAGAGCGCAGCATCTGGAGCA 745
Qy 830 AAGAGAGGATATAACAGAAAGCAGCCTCTC 861
Db 746 AAGAGAGGATATAACAGAAAGCAGCCTCTC 777
RESULT 12
BC023092
LOCUS
DEFINITION Mus musculus, Nedda 4, 1079 bp mRNA linear HTC 20-SEP-2002
ACCESSION BC023092 GI:18605637
VERSION BC023092.1
KEYWORDS HTC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 1079)
Straussberg, R.
Direct Submission
Submitted (04-FEB-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapb-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAK Plate: 36 Row: h Column: 6
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein
This clone has the following problem: frame shifted.
FEATURES
Location/Qualifiers
1..1079
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:3989996"
/tissue_type="Mammary tumor metastasized to lung. Tumor
arose spontaneously from a senescent normal mammary
(clonal) outgrowth infected with the virus "MTV."
/clone_lib="NCI_CGAP_Lu29"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
ORIGIN
Query Match 71.4%; Score 614.6; DB 11; Length 1079;
Best Local Similarity 85.3%; Pred. No. 4.2e-103;

Db	400	GGCGTGCACACTGCCACGCGCTCTTTGTTCCCCAGCATGGAGATCATCGGAGCTGGAGTTC	459
Qy	124	GTTCAGATCATCATCTCGTGTGTGATGATGGTGTGATGGTGTGATCATCGTGCCTG	183
Db	460	GTCAAAATCGTGTCTCATCTCGTGTGTGATGATGGTGTGATGGTGTGATGATACGTGCCTG	519
Qy	184	CTGAGCCACTACAAAGCTGTCTGCACGGTCTTCATCAGCCGCGACACACGAGGGGGGAGG	243
Db	520	CTGAGCCACTACAAAGCTGTCTGAGCCGCTCTTCATCAGCGACACACGAGCCAGGAGG	579
Qy	244	AGAGAAAGATCCCTGTCTCTCAGAAGATGCTGTGCGCTTCGAGAGCAAGTGTCAAGC	303
Db	580	AGAGACGATGGAATGTCTCTCGAAGATGCTCTGCGCTTCAGAGAGTACGGTGTCAAG-	638
Qy	304	AACGGAATCCAGAGCGCAGGCTTAGCGCCCGCTTCGCGCCACCGACCGCTGGCCGCTG	363
Db	639	--TGGAAATGCGGAGCCACAGGCTTAGTGCCTCCGCTTCGCGCCACTGACCGACTCGCTGTG	696
Qy	364	CCGCGCTTCGCCACGCGGGAGCGCTTCACCGCTTCGAGCCCACTATCCGTACCTGCAG	423
Db	697	CCCCCTTCATCAGCGG-----AGCCGATTCCAACCCACTACCCCTACCTGCAG	747
Qy	424	CAGAGATCAGCTGCCACCCACCATCTCGCTGTCAAGCGGGAGGAGCCCCACCCCTAC	483
Db	748	CACGAATTTGCCCTGCCCACCAACCATCTACTCTGTCTGATGGGAGGAGCCCCACCCCTAC	807
Qy	484	CAGGCGCCCTGCACCTTCCAGCTTCGGGACCCCGAGCAGCGTGGAACTGAAACCGGGAG	543
Db	808	CAGGCGCCCTGCACCTTCCAGCTACGGGACCCCTGAGCAACAGCTGGAGCTGAAACCGGGAA	867
Qy	544	TCGTTGGCGCACCCCCCAAACAGAACCATCTTCGACAGTCACTGTATGGATGTGCCAGG	603
Db	868	TCTGTGGCGCACCCCTTAACCGGACCATCTTCGACAGTCACTTATAGACACCATG	927
Qy	604	CTGGCGGCGCCCTCCCGCCCGACAGTAATTCGGGCATCAGGCGCAGTGTCTACGGGAGC	663
Db	928	CTGGGGGCGCCCTGTCCCGCCAGCAGTAATCTCGGGCATCAGCGCCACTCTCTACAGCAGC	987
Qy	664	GGCGGGCGCATGGAGGGGCGCGCCCACTACAGCGAGTTCATCGGCCACTACCCGGGG	723
Db	988	GGTGGGGCATGGAGGGGCGCGCCCACTACAGCGAGTTCATGGGCCACTACCTTGGC	1047
Qy	724	TCCTCCTTCCAGCACCGCAGCAGTGGGCGCGCTTCCTTGTCTGGAGGGGACCCGGCTC	783
Db	1048	TCCTCCTTCCAGCACCGCAAGTAACGGGGCATCTCTCCCTGTAGAGGGGACCCGGCTC	1107
Qy	784	CACCAACACACATCTGGCGCCCTAGAGAGCGACGCATCTTGGACCAAGAGAGGATAAA	843
Db	1108	CATCACTCGCACTTGCCCACTGGA-----GAACAGGAGGAGGAGAA	1152
Qy	844	CAGAAAGGACACCCCTCTC	861
Db	1153	CAGAAAGGTCAACCCCTC	1170

RESULT 15
BQ690750
LOCUS
DEFINITION
BQ690750
5', mRNA sequence.
ACCESSION
BQ690750
VERSION
BQ690750.1
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 890)
NH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE
Unpublished (1999)
JOURNAL
Contact: Robert Strausberg, Ph.D.
COMMENT
Email: cgapbs-remail.nih.gov

```

Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: the I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM2368 row: h column: 14
High quality sequence stop: 627.
Location/Qualifiers
1. .890
/organism="Homo sapiens"
/mol type="mRNA"
/db xref="taxon:9606"
/clone="IMAGE:6209341"
/tissue_type="ductal carcinoma, cell line"
/lab host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 110"
/notes="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)".
Note: this is a NIH MGC Library."

```

ORIGIN

Query Match	57.2%	Score 578.4	DB 13	Length 890
Best Local Similarity	53.5%	Fred. No. 1.9e-96		
Matches 615	Conservative 0	Mismatches 41	Indels 2	Gaps 1
QY	87	TTTGTTCCAGCATGGAGATCA	CGGAGCTGGAGTTTGTTCAGATCATCATCATCGTGGT	146
Db	105	TCCTCTCGAAACCAGGCAATGG	CGGAGCTGGAGTTTGTTCAGATCATCATCATCGTGGT	164
QY	147	GGTGATCATGTCATGTTGGTGTG	ATCATCGTGTCTGTAGCCACTACAAAGCTGTCTGC	206
Db	165	GGTGATCATGTCATGTTGGTGTG	ATCATCGTGTCTGTAGCCACTACAAAGCTGTCTGC	224
QY	207	ACGCTCTTTCATCAGCCGGCACA	CGCCACGAGGGCGGAGAGAGAGATGCCCTGTCTCAGA	266
Db	225	ACGCTCTTTCATCAGCCGGCACA	CGCCACGAGGGCGGAGAGAGAGATGCCCTGTCTCAGA	284
QY	267	AGGATGCTGTGGCCCTCGGAGAG	CACAGTGTCTAGGCAACGGAAATCCAGAGCCCGCAGGT	326
Db	285	AGGATGCTGTGGCCCTCGGAGAG	CACAGTGTCTAGGCAACGGAAATCCAGAGCCCGCAGGT	344
QY	327	CTACGCCCGCGCTCGGCCAC	CCGACGCCCTGGCCGCTTCGCCACAGCGGGAGCG	386
Db	345	CTACGCCCGCGCTCGGCCAC	CCGACGCCCTGGCCGCTTCGCCACAGCGGGAGCG	404
QY	387	CTTCCACCGCTTCAGCCGCA	CCCTATCCGTCAGATCGAGATCGACCTGCCACCCAC	446
Db	405	CTTCCACCGCTTCAGCCGCA	CCCTATCCGTCAGATCGAGATCGACCTGCCACCCAC	464
QY	447	CATCTCGCTGCAGACGGGAGG	AGCCCCACCTTACCAGGGCCCCCTGCACCTCTCCAGCT	506
Db	465	CATCTCGCTGCAGACGGGAGG	AGCCCCACCTTACCAGGGCCCCCTGCACCTCTCCAGCT	524
QY	507	TCGGGAACCCGAGCAGCAGCT	GGAACTGAAACCGGAGTTCGGTGGCGCACCCGCCAAACAG	566
Db	525	TCGGGAACCCGAGCAGCAGCT	GGAACTGAAACCGGAGTTCGGTGGCGCACCCGCCAAACAG	584
QY	567	AACCATCTTCACACAGTACCT	GATGGATAGTCCAGGCTGGGCGGCCCTGTGCCCCCCCAG	626
Db	585	AACCATCTTCACACAGTACCT	GATGGATAGTCCAGGCTGGGCGGCCCTGTGCCCCCCCAG	644
QY	627	CAGTAACTCGGCATCAGCGCA	CGTGTCTAGGCAGCGCGCGGCGCATGGAA--GGGGCGCG	684
Db	645	CAGTAACTCGGCATCAGCGCA	CGTGTCTAGGCAGCGCGCGGCGCATGGAA--GGGGCGCG	704
QY	685	CGGCCCACTCAGCGGAGGT	CATCGGCCACTACCGGGGTCTCTCTTCACACACACAGC	742

Db 705 CGCCCCACCTACAGGCAAGGCATCTGGCCACCTACCCGGGGTCCCTCCTTTTCAGC 762

Search completed: May 12, 2004, 05:18:16
JOB time : 2593.37 secs

ALIGNMENTS

```

CORRECTION FILING DATE: 2000-01-28
PRIORITY APPLICATION NUMBER: 60/178,772
PRIORITY FILING DATE: 2000-01-28
PRIORITY APPLICATION NUMBER: 60/179,045
PRIORITY FILING DATE: 2000-01-31
NUMBER OF SEQ ID NOS: 67
SOFTWARE: Patent Ver. 2.1

```

```
Query Match      87.5%; Score 753.6; DB 4; Length 1140;
Best Local Similarity 99.5%; Pred. No. 9e-160;
Matches 756; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

102	QY	GGAGATCAGGAGCTGGAGCTTTGTCAGATCATCATCTGCTGGTGATGATGGTGAT	161
91	DB	GGCAATGGCGGAGCTGGAGTTTGTTCAGATCATCATCTGCTGGTGATGATGGTGAT	150
162	QY	GGTGGTGGTGAATCAGCTGCCTCTGAGCGAATPACAAGCTGTCTGCACGGTCTTTCATCAG	221
151	DB	GGTGGTGGTGAATCAGCTGCCTCTGAGCGAATPACAAGCTGTCTGCACGGTCTTTCATCAG	210
222	QY	CCGGCACAGCCAGGGCGGAGAGAGAGATGCCCTGTCTCACAAGGATGCCCTGTGGCC	281
211	DB	CCGGCACAGCCAGGGCGGAGAGAGAGATGCCCTGTCTCACAAGGATGCCCTGTGGCC	270
282	QY	CTCGAGAGCACAGTGTCAAGGCACGGAAATCCAGAGCCGAGGCTACGCCCGCGCTCG	341
271	DB	CTCGAGAGCACAGTGTCAAGGCACGGAAATCCAGAGCCGAGGCTACGCCCGCGCTCG	330
342	QY	GCCACCGACCGCTGGCGGTGCGCGCTTTCGCCCCAGGGGAGGCGCTTCCACCGCTTCCA	401

SIMMARTES

Result No.	Score	Query			ID	Description
		Match	Length	DB		
1	753.6	87.5	1140	4	US-09-769-482-1	Sequence 1, Appli
2	752.8	87.4	759	4	US-09-769-482-2	Sequence 2, Appli
3	352.2	40.9	921	4	US-09-091-952A-7	Sequence 7, Appli
4	352.2	40.9	8065	4	US-09-091-952A-6	Sequence 6, Appli
5	284.2	33.0	867	4	US-09-091-952A-8	Sequence 8, Appli
6	90.4	10.5	391	4	US-09-621-976-2162	Sequence 2162, Ap
7	48	5.6	696	4	US-09-252-991A-13687	Sequence 13687, A
8	48	5.6	1452	4	US-09-252-991A-13650	Sequence 13650, A
9	47.8	5.6	8438	1	US-07-945-283-1	Sequence 1, Appli
10	47.4	5.5	1215	4	US-09-266-965-42	Sequence 42, Appli
11	47.4	5.5	53500	4	US-09-266-965-76	Sequence 76, Appli
12	46.2	5.4	561	4	US-09-252-991A-4586	Sequence 4586, Ap
13	46.2	5.4	1203	4	US-09-252-991A-4391	Sequence 4391, Ap
14	46.2	5.4	1611	4	US-09-252-991A-4200	Sequence 4200, Ap
15	46.2	5.4	15378	3	US-08-785-430-1	Sequence 1, Appli
16	46	5.3	3756	4	US-09-252-991A-692	Sequence 692, App
17	46	5.3	3774	4	US-09-252-991A-719	Sequence 719, App
18	46	5.3	6885	4	US-09-252-991A-660	Sequence 660, App
19	45.6	5.3	1119	4	US-09-252-991A-7780	Sequence 7780, Ap
20	45.6	5.3	1887	4	US-09-252-991A-7920	Sequence 7920, Ap
21	45.4	5.3	424	2	US-08-476-176B-7	Sequence 7, Appli
22	45.4	5.3	424	3	US-08-127-721A-7	Sequence 7, Appli
23	45.4	5.3	424	3	US-08-485-246A-7	Sequence 7, Appli
24	44	5.1	1293	2	US-08-924-440-1	Sequence 1, Appli
25	44	5.1	44377	2	US-08-804-227C-7	Sequence 7, Appli
26	44	5.1	44377	2	US-08-804-198-1	Sequence 1, Appli
27	44	5.1	4403765	3	US-09-103-840A-2	Sequence 2, Appli

SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/091,952A
FILING DATE: 19-Apr-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/029,278
FILING DATE: 28-OCT-1996
APPLICATION NUMBER: PCT/US97/19381
FILING DATE: 28-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Timothy L.
REGISTRATION NUMBER: 35,367
REFERENCE/DOCKET NUMBER: 015280-297100US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 921 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1...921
OTHER INFORMATION: Clone 22 coding region
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-091-952A-7

Query Match 40.9%; Score 352.2; DB 4; Length 921;
Best Local Similarity 73.1%; Pred. No. 5.1e-70;
Matches 482; Conservative 0; Mismatches 168; Indels 9; Gaps 2;
QY 94 CAGAGCATGGAGATCAGGAGCTGGAGTTGTTTCAGATCATCATCTGTTGTTGATG 153
DB 166 CCGGGCATCTTCACTCGGAGTGGAGTTGCGCCAAATCATCATCTGTTGTTGATG 225
QY 154 ATGGTATGGTGGTGGTATCAGTGCCTGCTGAGCCACTCAAGTGTCTCAGCGTCC 213
DB 226 ACGGTATGGTGGTGGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 285
QY 214 TTCTCAGCGGGCAGACAGCGGGGGGAGAGAGATGCCCTGTCTCAGAGGATGC 273
DB 286 TTCTCAACCGCCGCAACAGAGCGGAGCGGGAGCGGCTGCCGAGGAGGATGC 345
QY 274 CTGTGGCCCTCGGAGAGACAGTGTCTCAGGCAACGGAATCCAGAGCGGAGTCTACGCC 333
DB 346 CTGTGGCCCTCGGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 399
QY 334 CCGCTCGGCGCCACCGACCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTG 393
DB 400 GCGCCGCGTCCAGGAGCAGAGTTACAGCGCGCTCTTCATCCAGAGGATCGCTTCAGC 459
QY 394 CGCTTCAGCGCCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 453
DB 460 CGCTTCAGCGCCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 519
QY 454 CTGTCAAGCGGGAGAGCGCCCGCCCTACCGGCGCCCTGCGCGCTGCGCGCTGCGCGCTG 513
DB 520 CTGTCCGAGCGTGAAGAGCAGCTCTTACCGGCGCCCTGCGCGCTGCGCGCTGCGCGCTG 579
QY 514 CCGGAGCAGCGTGAAGTGAACCGGAGTGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCG 573
DB 580 CCGTGAACAGCAGATGAATCAACCGAGAGTCCGTGAGGGGCGCCCGCCCGCCCGCCCG 639
QY 574 TTGACAGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 630
DB 640 TTTGACAGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 699
QY 631 AACTCGGGCATCAGCGCCAGCTGCTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 690

DB 700 AACTCGGGCATCAGTCAAGCAGCTTCAGATGTAACGGAGGATGAGGGGCCACCCCC 759
QY 691 ACCTACAGCGAGGTCTATCGGCCACTACCGGGGTCTCTTCAGACCGAGAGCAG 749
DB 760 ACATACAGCGAGGTGATGGGCCACCAACCGAGGGGCTCTTCTCCATCACCAGCGCAG 818
RESULT 4
US-09-091-952A-6
Sequence 6, Application US/09091952A
Patent No. 6458532
GENERAL INFORMATION:
APPLICANT: Detera-Wadleigh, Sevilla D.
Gershon, Elliot S.
Badner, Judith A.
Goldin, Lynn R.
Berrettini, Wade H.
Yoshikawa, Takeo
Sanders, Alan R.
Esterling, Lisa E.
TITLE OF INVENTION: Chromosomal Markers and Diagnostic
NUMBER OF SEQUENCES: 197
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/091,952A
FILING DATE: 19-Apr-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/029,278
FILING DATE: 28-OCT-1996
APPLICATION NUMBER: PCT/US97/19381
FILING DATE: 28-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Timothy L.
REGISTRATION NUMBER: 35,367
REFERENCE/DOCKET NUMBER: 015280-297100US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 8065 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: -
LOCATION: 1...8065
OTHER INFORMATION: Clone 22
FEATURE:
NAME/KEY: CDS
LOCATION: 116...1036
OTHER INFORMATION: Clone 22 coding region
FEATURE:
NAME/KEY: misc feature
LOCATION: 452...505
OTHER INFORMATION: alternatively spliced portion
FEATURE:

RESULT 11

US-09-266-965-76
; Sequence 76, Application US/09256965
; Patent No. 6493348
; GENERAL INFORMATION:
; APPLICANT: Sherman, D
; APPLICANT: Mao, Y
; APPLICANT: Varoglu, M
; APPLICANT: He, M
; APPLICANT: Sheldon, P
; TITLE OF INVENTION: Mitomycin biosynthetic gene cluster
; FILE REFERENCE: 600.456051
; CURRENT APPLICATION NUMBER: US/09/266,965
; CURRENT FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: US 08/624,447
; EARLIER FILING DATE: 1996-08-19
; EARLIER APPLICATION NUMBER: PCT/US94/11279
; EARLIER FILING DATE: 1994-10-05
; EARLIER APPLICATION NUMBER: US 08/133,963
; EARLIER FILING DATE: 1993-10-07
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 76
; LENGTH: 53500
; TYPE: DNA
; ORGANISM: Streptomyces lavendulae
US-09-266-965-76

Query Match 5.5%; Score 47.4; DB 4; Length 53500;
Best Local Similarity 47.2%; Pred. No. 0.2;
Matches 176; Conservative 0; Mismatches 196; Indels 1; Gaps 1;
QY 446 CCAATCTGCTGTCAGACGGGAGAGCCGCCACCTTACACAGGCGCCCTCGACCTCCAGC 505
Db 13517 CCACGTGGTGTGCGGTGACAGGAAGCGCGCGGTGCTCTCCACCGTCTCCCGCGGC 13576
QY 506 TTCGGACCCCGACGACCTGGAATGAACCGGAGTCGGTGGCGGACACCCCAACA 565
Db 13577 CTGGCCGCGCGTGGCGGAGAGCGCGGTCCAGTCGCGCTGGCCCGGCTCA 13636
QY 566 GAACCATCTTCGACAGTACCTGATGATAGTCCAGGTGGCGGCGCCCTGCCCCCA 625
Db 13637 TCCAGTCCCGGACATGCTGCTCGCGCCCGGACACCGCGGCCACCCAGCAGTGCA 13696
QY 626 GCAGTAATCTGGCATCAGCGCCACGTCGTACGCGAGCGCGGCGCATGGAGGCGCC 685
Db 13697 GCGCCACGTCAGAGTGCCACGCCAGTGTGAGACGCGCGCGCCCGGAGCGCGC 13756
QY 686 CGCCACCTACAGCGAGGTGATCGGCCACTACCGGGGTCTCTTCCAGCACCGACAGA 745
Db 13757 GGTGACGAACCAACCGGTGCGGTGGGATGCGGTGGCGCGGATCCAGC-TCAGCCG 13815
QY 746 GCAGTGGCGCCCTCTTCTGAGGGGACCGGCTCCACACACAGACATCGCGCC 805
Db 13816 AACTGCGACGTTGCCCGACGAGGAGCAGCTCCGCCAGGCGGCGAGACATCGGTGG 13875
QY 806 TAGAGAGCGCAGC 818
Db 13876 TGCCGCGCGCGC 13888

RESULT 12

US-09-252-991A-4586
; Sequence 4586, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4586
; LENGTH: 561
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4586

Query Match 5.4%; Score 46.2; DB 4; Length 561;
Best Local Similarity 51.2%; Pred. No. 0.12;
Matches 108; Conservative 0; Mismatches 103; Indels 0; Gaps 0;
QY 417 CCTGCAGCAGCAGATCGACCTGGCCACCCACCATCTCGCTGTCAGACGGGAGAGCCCC 476
Db 175 CCTGCTGACGATATCGGCCACCTCTAGAGAGACCCGGCGCAGATCCACGAGGAGACCT 234
QY 477 ACCCTACAGGCGCCCTGCAACCTTCAGCTTCGGGACCCCGAGCAGCAGCTGGAACTGAA 536
Db 235 GCGCCACGAGGAATTCGGCGCAGCGCTGCTGCGGAGCTGTTCCAGAGTCGGTCTGCA 294
QY 537 CGGGAGTCGGTGGCGGCGACCCCAACAGAACCATCTTCACAGTGCACCTGATGATAG 596
Db 295 GCCGTGGCTTCGACGATTCGCCAAGCGCTTCTCTGCGGTGGACCCGAGCTACCA 354
QY 597 TCCAGGCTGGCGCGCGCCCTGCTCCCGCCCGCAGC 627
Db 355 CGCCAGCCTGTGCGCGGCTCCCGGCACAGC 385

RESULT 13

US-09-252-991A-4391
; Sequence 4391, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4391
; LENGTH: 1203
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4391

Query Match 5.4%; Score 46.2; DB 4; Length 1203;
Best Local Similarity 51.2%; Pred. No. 0.15;
Matches 108; Conservative 0; Mismatches 103; Indels 0; Gaps 0;
QY 417 CCTGCAGCAGCAGATCGACCTGGCCACCCACCATCTCGCTGTCAGACGGGAGAGCCCC 476
Db 783 CCTGCTGACGATATCGGCCACCTCTAGAGAGACCCGGCGCAGATCCACGAGGAGACCT 842
QY 477 ACCCTACAGGCGCCCTGCAACCTTCAGCTTCGGGACCCCGAGCAGCAGCTGGAACTGAA 536
Db 843 GCGCCACGAGGAATTCGGCGCAGCGCTGCTGCGGAGCTGTTCCAGAGTCGGTCTGCA 902
QY 537 CGGGAGTCGGTGGCGGCGACCCCAACAGAACCATCTTCACAGTGCACCTGATGATAG 596
Db 903 GCCGTGGCTGACGATTCGGCAAGCGCTTCTCTGCGGTGGACCCGAGCTACCA 962
QY 597 TCCAGGCTGGCGCGCGCCCTGCTCCCGCCCGCAGC 627
Db 963 CGCCAGCCTGTGCGCGGCGCTCCCGGCACAGC 993

RESULT 15
US-08-785-420-1/c
; Sequence 1, Application US/08785420
; Patent No. 6001976
; GENERAL INFORMATION:
; APPLICANT: MacLennan, David H
; APPLICANT: O'Brien, Peter J
; TITLE OF INVENTION: DIAGNOSIS FOR PORCINE MALIGNANT
; TITLE OF INVENTION: HYPERTHERMIA
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bell, Seltzer, Park & Gibson
; STREET: P.O. Drawer 34009
; CITY: Charlotte,
; STATE: No. 6001976th Carolina 28234
; COUNTRY: U.S.A.
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/785,420
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/249,388

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	861	100.0	861	6	AX392419	Sequence
2	861	100.0	1321	6	AX392417	Sequence
3	861	100.0	4839	9	AF305616	Homo sapi
4	859.4	99.8	1383	6	AX775889	Sequence
5	839.4	97.5	1061	9	BC015918	Homo sapi
6	755.2	87.7	969	6	BD272494	Secrated
7	754.2	87.6	1141	9	AF224278	Homo sapi
8	753.8	87.5	1913	6	BD272544	Secrated
9	753.6	87.5	969	6	BD272514	Secrated
10	753.6	87.5	969	6	BD272515	Secrated
11	753.6	87.5	969	6	BD272516	Secrated
12	753.6	87.5	1140	6	AR336830	Sequence
13	752.8	87.4	759	6	AR336831	Sequence
14	752.6	87.4	1085	6	AX775887	Sequence
15	752.4	87.4	759	6	BD272545	Secrated
16	752	87.3	1818	9	AY128643	Homo sapi
17	751.2	87.2	756	6	BD272495	Secrated
18	750.4	87.2	753	6	BD272534	Secrated
19	614.6	71.4	1379	10	BC036995	Mus muscu
20	612.4	71.1	878	6	AX392428	Sequence
21	546.4	63.5	61505	9	AF305426	Homo sapi
22	546.4	63.5	130435	9	HS718J7	Human DNA
23	544.4	63.2	1583	6	AX593655	Sequence
24	462.2	53.7	1713	6	BD272504	Secrated
25	461.2	53.6	651	6	BD272505	Secrated
26	460.6	53.5	1713	6	BD272517	Secrated
27	460.6	53.5	1713	6	BD272518	Secrated
28	460.6	53.5	1713	6	BD272519	Secrated
29	458.2	53.2	648	6	BD272535	Secrated
30	445.6	51.8	693	6	AX392430	Sequence
31	421	48.9	651	10	AF220208	Mus muscu
32	417.4	48.5	812	6	AX011709	Sequence
33	417.4	48.5	812	6	BD26320	Pancrreati
34	401.4	46.6	408	6	AX071267	Sequence
35	400.6	46.5	673	6	AX525744	Sequence
36	373.2	43.3	156698	10	AL837509	Mouse DNA
37	373.2	43.3	175754	2	AC110189	Mus muscu
38	373.2	43.3	231930	2	AC134911	Mus muscu
39	353.6	41.1	249554	2	AC139417	Rattus no
40	353.6	41.1	258632	2	AC111878	Rattus no
41	352.2	40.9	921	6	AR233384	Sequence
42	352.2	40.9	8065	6	AR233383	Sequence
43	352.2	40.9	8494	9	AF009424	Homo sapi
44	350	40.7	2170	6	AX713513	Sequence
45	350	40.7	2170	9	AK055028	Homo sapi

ALIGNMENTS

RESULT 1	AX392419	AX392419	861 bp	DNA	linear	PAT 23-MAR-2002
LOCUS	AX392419	Sequence 3 from Patent WO0216416.				
DEFINITION	AX392419	Sequence 3 from Patent WO0216416.				
ACCESSION	AX392419					
VERSION	AX392419.1	GI:19700734				
KEYWORDS						
SOURCE		Homo sapiens (human)				
ORGANISM		Homo sapiens				
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
		Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo				
REFERENCE	1					
AUTHORS		Lee, R.T., Landschulz, K.T., Kennedy, S.P., Thompson, J.F. and Turi, T.G.				
TITLE		Diagnosis and treatment of cardiovascular conditions				

Pred. No. is the number of results predicted by chance to have a

OM nucleic - nucleic search, using sw model

Run on: May 11, 2004, 21:10:37 ; Search time 3552.51 Seconds (without alignments)
10504.776 Million cell updates/sec

Title: US-09-934-249-3
Perfect score: 861
Sequence: 1 atgcacgccttgatgggggt.....aacagaaagacacccctctc 861

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rtd.*
36: em.htg.nam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

	QY	421	CAGCAGGATCGACTGTGCCACCCACCATCTCGCTGTAGACGGGGAGAGAGCCCCCACCC	480
	Db	741	CAGCAGGATGACCTTCGGGCCACCATCTCGCTGTAGACGGGGAGAGAGCCCCCACCC	800
	QY	481	TACCAGGGCCCTTGCAACCTTCCAGCTTTGGGACCCCAGCAGCAGCTGGAACTGAACCGG	540
	Db	801	TACCAGGGCCCTTGCAACCTTCCAGCTTTGGGACCCCAGCAGCAGCTGGAACTGAACCGG	860
	QY	541	GAGTCGGTGGGCACCCCCNACAGAACATCTTCGACAGTCACTGTATGGATATGTC	600
	Db	861	GAGTCGGTGGGCACCCCCNACAGAACATCTTCGACAGTCACTGTATGGATATGTC	920
	QY	601	AGGCTGGGGCCCTTGCCTCCCGCCAGCACTTCGGGATCAGCGCAGCTGTACGGC	660
	Db	921	AGGCTGGGGCCCTTGCCTCCCGCCAGCACTTCGGGATCAGCGCAGCTGTACGGC	980
	QY	661	AGCGGGGGCGATGAGGGGGCGCGCCAGCTTACAGCGAGGTCTATCGGCCACTACCGG	720
	Db	981	AGCGGGGGCGATGAGGGGGCGCGCCAGCTTACAGCGAGGTCTATCGGCCACTACCGG	780
	QY	721	GGTCTCTCTTCAGCACCAGCAGAGCAGTGGCGCCCTCTCTTCTGTGAGGGGACCCCG	780
	Db	1041	GGTCTCTCTTCAGCACCAGCAGAGCAGTGGCGCCCTCTCTTCTGTGAGGGGACCCCG	1100
	QY	781	CTCCACACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAAGAGAGAT	840
	Db	1101	CTCCACACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAAGAGAGAT	1160
	QY	841	AAACAGAAAGCAGCCTCTCTC	861
	Db	1161	AAACAGAAAGCAGCCTCTCTC	1181
	RESULT 5			
	BC015918		1061 bp	linear
	LOCUS			mRNA
	DEFINITION			Homo sapiens transmembrane, prostate androgen induced RNA, mRNA
			(CDNA clone MGC:20374 IMAGE:4559576), complete cds.	
	ACCESSION			BC015918
	VERSION			BC015918.1
	KEYWORDS			MGC.
	SOURCE			Homo sapiens (human)
	ORGANISM			Homo sapiens
	REFERENCE			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	AUTHORS			Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
				1 (bases 1 to 1061)
				Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
				Klausner,R.D., Collins,F.S., Wagner,K.H., Shenmen,C.M., Schuler,G.D.,
				Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
				Hopkins,R.F., Jordan,H., Moore,T., Max,A.I., Wang,J., Hsieh,F.,
				Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
				Stapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L.,
				Scheetz,T.E., Brownstein,M.J., Uedin,T.S., Lasham,A., Venter,
				Carninci,P., Prange,C., Raja,S.S., Loquellano,N.A., Peters,G.J.,
				Abrahamson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,
				McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
				Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
				Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
				Fahy,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
				Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
				Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
				Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
				Butterfield,Y.S., Krzywinska,M.I., Skalska,U., Smalley,D.E.,
				Schnierch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
				Generation and initial analysis of more than 15,000 full-length
				human and mouse cDNA sequences
				Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
	JOURNAL			2238457
	MEDLINE			12477932
	PUBMED			2
	REFERENCE			(bases 1 to 1061)
	AUTHORS			Strausberg,R.
	TITLE			Direct Submission

Submitted (15-OCT-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs@mail.nih.gov
Tissue Procurement: DCTB/DTP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzyzinski, Rea Kutsche, Oliver Lee, Soo
Lee, Victor Ling, Carrie Mathewson, Candice Mcueavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Teai, Natasha van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: ITAL Plate: 29 Row: e Column: 5
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

FEATURES	
SOURCE	

```
1. 1961
/organism="Homo sapiens"
/mcl_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:20374 IMAGE:4559376"
/tissue_type="Kidney, renal cell adenocarcinoma"
/clone_lib="NIH MGC.14"
/lab_host="DH10B-R"
/notes="Vector: pOTB7"
```

gene

```
1. .1061
/gene="TMEPAI"
/note="synonym: PMEPAI"
/db_xref="LocusID:56937"
/db_xref="MIM:606564"
131. .844
```

SDS

```

/codon_start=1
/product="TM6PAI protein"
/protein_id="AAH15918.1"
/db_xref="GI:16198475"
/db_xref="LocusID:56937"
/translation="MMVMVVITCL
SESTVSGNGIPEQVYAPPTDRLL
SDGEERPPYGGPCTQLRDPEQGLE
LSAATWSYKSGDKMGEPHPYSEV
LESAATWSYKSGDKMGHPPL"

```

ORIGIN

Query Match 97.5%; Score 839.4; DB 9; Length 1061;
Best Local Similarity 99.9%; Pred. No. 1.8e-133;
Matches 840; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	21	CAACAGCACCGCGCCGCGCCGCCGGGCAGCCCCAATGTCTCTGTGAAGTGAACACTGCAA	80
Db	1	CAACAGCACCGCGCCGCGCCGCCGGGCAGCCCCAATGTCTCTGTGAAGTGAACACTGCAA	60

Qy	81	ACGCTCTTTTCCAGAGCATGGAGATCA	CGGAGCTGGAGTTTGTTCAGATCATCAT	140
Db	61	ACGCTCTTTTTCAGAGCATGGAGATCA <td>CGGAGCTGGAGTTTGTTCAGATCATCAT</td> <td>120</td>	CGGAGCTGGAGTTTGTTCAGATCATCAT	120

QY	141	CGTGGTGGTGATGATGCTGCTGGTGGTGATCACGTGCTGCTGAGCCACTACAGCT	200
Dp	121	CGTGGTGGTGATGATGCTGCTGGTGGTGATCACGTGCTGCTGAGCCACTACAGCT	180

201 GTCTGCACGGCTCTTTCATCAGCGGCGACGCCAGGGCGGAGGACAGAGATGCCCTGTC 260

[illegible]

PC GOIN33/50,GOIN33/53,GOIN33/53,GOIN33/566,CL2N15/00,CL2N15/00 CC
Secreted proteins and nucleic acids encoding them FH Key
FT CDS Location/Qualifiers
source 1..969
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source
Query Match 87.7%; Score 755.2; DB 6; Length 969;
Best Local Similarity 99.6%; Pred. No. 4.1e-119;
Matches 757; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

ORIGIN
102 GGAGATCACGAGCTGGAGTTTGTTCAGATCATCATCGTGGTGTGATGATGAT 161
Db 2 GGAGATCGCGAGCTGGAGTTTGTTCAGATCATCATCGTGGTGTGATGATGAT 61
QY 162 GGTGGTGGTATCAGCTGCTGTGAGCCACTACAGCTGTCTGCACGCTCTTCATCAG 221
Db 62 GGTGGTGGTATCAGCTGCTGTGAGCCACTACAGCTGTCTGCACGCTCTTCATCAG 121
QY 222 CCGGCACAGCAGCGGGCGGAGGAGAGATGCCCTGTCTCAGAAAGGATGCTGTGGCC 281
Db 122 CCGGCACAGCAGCGGGCGGAGGAGAGATGCCCTGTCTCAGAAAGGATGCTGTGGCC 181
QY 282 CTGGAGACACAGTGTGAGCAACGGAATCCAGAGCCGAGGTCTAGCCCGGCTCG 341
Db 182 CTGGAGACACAGTGTGAGCAACGGAATCCAGAGCCGAGGTCTAGCCCGGCTCG 241
QY 342 GCCACACAGCGCTGCGGCTGCGGCTTGCAGCGGAGCGCTTCCACGCTTCA 401
Db 242 GCCACACAGCGCTGCGGCTTGCAGCGGAGCGCTTCCACGCTTCA 301
QY 402 GCCACCTATCGTACTGTGAGCAGCAGATCGACCTGCGACACCATCTCGTGTGAGA 461
Db 302 GCCACCTATCGTACTGTGAGCAGCAGATCGACCTGCGGCGGAGGATCTCGTGTGAGA 361
QY 462 CGGGAGAGCGCCACCTACACAGGCGCCCTGACCTCCAGCTTCGGAGCCCGGAGCA 521
Db 362 CGGGAGAGCGCCACCTACACAGGCGCCCTGACCTCCAGCTTCGGAGCCCGGAGCA 421
QY 522 GCAGCTGGAACTGAACCGGAGTCGTGCGGCGACCCGCCAAACAGAACCATCTTCAGACAG 581
Db 422 GCAGCTGGAACTGAACCGGAGTCGTGCGGCGACCCGCCAAACAGAACCATCTTCAGACAG 481
QY 582 TGACCTGATGATAGTSCCAGGCTGGGCGCCCTGCGCCCGCCAGCAAGTAACCTCGGCGCAT 641
Db 482 TGACCTGATGATAGTSCCAGGCTGGGCGCCCTGCGCCCGCCAGCAAGTAACCTCGGCGCAT 541
QY 642 CAGCGCCACGTGCTACCGGACGCGCGCGCGATGGAGGGCGCGCCGCCACCTTACAGCGCA 701
Db 542 CAGCGCCACGTGCTACCGGACGCGCGCGCGATGGAGGGCGCGCGCCGCCACCTTACAGCGCA 601
QY 702 GGTATCGGCACTACCGGCGTCTCTTCCAGCACCCAGCAGCAGTGGCGCGCCCTC 761
Db 602 GGTATCGGCACTACCGGCGTCTCTTCCAGCACCCAGCAGCAGTGGCGCGCCCTC 661
QY 762 CTTGCTGGAGGGGACCGGCTTCCACCAACATCGCGCCCTTAGAGAGCGCAGCCAT 821
Db 662 CTTGCTGGAGGGGACCGGCTTCCACCAACATCGCGCCCTTAGAGAGCGCAGCCAT 721
QY 822 CTGGAGCAAGAGAGGATTAACAGAGGACACCCCTTC 861
Db 722 CTGGAGCAAGAGAGGATTAACAGAGGACACCCCTTC 761

RESULT 7
AF224278 1141 bp mRNA linear PRI 18-JUL-2000
LOCUS Homo sapiens PMEPAL protein (PMEPAL) mRNA, complete cds.
DEFINITION AF224278
ACCESSION

AF224278.1 GI:9255808
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1141)
XU,L.D., Shanmugam,N., Segawa,T., Sesterhenn,I.A., McLeod,D.G.,
Moul,J.W. and Srivastava,S.
A novel androgen-regulated gene, PMEPAL, located on chromosome
20q13 exhibits high level expression in prostate
Genomics 66 (3), 257-263 (2000)
20334621
10873380
2 (bases 1 to 1141)
XU,L.D., Shanmugam,N., Segawa,T., Sesterhenn,I.A., McLeod,D.G.,
Moul,J.W. and Srivastava,S.
Direct Submission
Submitted (12-JAN-2000) CPDR, USUHS, 1530 East Jefferson Street,
Rockville, MD 20852, USA
Location/Qualifiers
1..1141
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20q13.31-13.33"
/cell_line="LNCaP"
1..1141
/gene="PMEPAL"
96..854
/note="type 1b transmembrane protein; expression is
induced in response to the synthetic androgen, R1881;
expression is abundant in, and restricted to, prostate
glandular epithelial cells; similar to the predicted
protein encoded by sequence deposited at GenBank Accession
Number NP_004329"
/codon_start=1
/product="PMEPAL protein"
/protein_id="AAF86322.1"
/db_xref="GI:9255809"
/translation="MALEFVQIIIVVMVMVITLLSHYKLSARGFISRHSG
RRRDALSSGCLWPSSESTVGNGIPPEOVYAPRPTDRDLAVPEAFQREPHRFOPTY
PYLQHEIDLPTLSLDEBPPYQGGCTQLQADPEQOLELNRESVAPNNRTIEDSD
LMDSGRLGGPCPPSSNGISATCYGSGRMEGPPPTYSIVIGTFGSGFQHQSSGPP
SLLEGTRHHTHIAPLESAATWSKEDKQKHPL"

ORIGIN
Query Match 87.6%; Score 754.2; DB 9; Length 1141;
Best Local Similarity 98.3%; Pred. No. 5.9e-119;
Matches 762; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 87 TTGTGTCAGAGCATGGAGATCAGGAGCTGGAGTTTGTTCAGATCATCATCGTGT 146
Db 77 TCTCTCGAAAACCAAGCAATGGCGAGCTGGAGTTTGTTCAGATCATCATCGTGT 136
QY 147 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 206
Db 137 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 196
QY 207 ACGGTCCTTCATCAGCCGSCACAGCCAGGCGCGGAGGAGAGAAGATGCCCTG 266
Db 197 ACGGTCCTTCATCAGCCGSCACAGCCAGGCGCGGAGGAGAGAAGATGCCCTG 256
QY 267 AGGATGCTGTGGCCCTCGGAGAGACAGTGTGAGCAACGGATCCAGAGCGCAG 326
Db 257 AGGATGCTGTGGCCCTCGGAGAGACAGTGTGAGCAACGGATCCAGAGCGCAG 316
QY 327 CTACGCCCCGCTCGGCCCCACCGACCGCTGCGCGCTGCGCGCTTCGCCACGCG 386
Db 317 CTACGCCCCGCTCGGCCCCACCGACCGCTGCGCGCTGCGCGCTTCGCCACGCG 376

[illegible]


```

PF 01-MAR-2000 JP 2000602247
PR 01-MAR-1999 US 60/122458
PI THOMAS M BARNES, DOUGLAS A HOLTZMAN, JOHN D SHARP, CHRISTOPHER C
PI FRASER
PC C12N15/09, C07K14/47, C07K16/18, C12N5/10, C12P21/02, C12Q1/68, PC
G01N33/15
PC G01N33/50, G01N33/53, G01N33/566, C12N15/00, C12N5/00 CC
Secreted proteins and nucleic acids encoding them FH Key
Location/Qualifiers
FT CDS Location/Qualifiers
(6)..(761).
FEATURES
Source
1..969
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 87.5%; Score 753.6; DB 6; Length 969;
Best Local Similarity 99.5%; Pred. No. 7.6e-119;
Matches 756; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 102 GGAGATCAGGAGTGGAGTTGTTTCAGATCATCATCATCGTGGTGGTGGTGGTAT 161
DB 2 GGAGATCAGGAGTGGAGTTGTTTCAGATCATCATCATCGTGGTGGTGGTGGTAT 61
QY 162 GGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 221
DB 62 GGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 121
QY 222 CCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 281
DB 122 CCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 181
QY 282 CTCGGAGAGCAGTGTGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 341
DB 182 CTCGGAGAGCAGTGTGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 241
QY 342 GCCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 401
DB 242 GCCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 301
QY 402 GCCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 461
DB 302 GCCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 361
QY 462 CTTGCTGGAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 821
DB 662 CTTGCTGGAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 721
QY 822 CTGGAGCAAGAGAGAGATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 861
DB 722 CTGGAGCAAGAGAGAGATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 761

```

```

RESULT 10
LOCUS BD272515
DEFINITION Secreted proteins and nucleic acids encoding them.
ACCESSION BD272515
VERSION BD272515.1 GI:33082283
KEYWORDS JP 2002539773-A/24.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 969)
Barnes, T.M., Holtzman, D.A., Sharp, J.D. and Fraser, C.C.
Secreted proteins and nucleic acids encoding them
Patent: JP 2002539773-A/24; 26-NOV-2002;
MILLENNIUM PHARMACEUTICALS INC
OS Homo sapiens (human)
PN JP 2002539773-A/24
PD 26-NOV-2002
PF 01-MAR-2000 JP 2000602247
PR 01-MAR-1999 US 60/122458
PI THOMAS M BARNES, DOUGLAS A HOLTZMAN, JOHN D SHARP, CHRISTOPHER C
PI FRASER
PC C12N15/09, C07K14/47, C07K16/18, C12N5/10, C12P21/02, C12Q1/68, PC
G01N33/15
PC G01N33/50, G01N33/53, G01N33/566, C12N15/00, C12N5/00 CC
Secreted proteins and nucleic acids encoding them FH Key
Location/Qualifiers
FT CDS Location/Qualifiers
(6)..(761).
FEATURES
Source
1..969
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 87.5%; Score 753.6; DB 6; Length 969;
Best Local Similarity 99.5%; Pred. No. 7.6e-119;
Matches 756; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 102 GGAGATCAGGAGTGGAGTTGTTTCAGATCATCATCATCGTGGTGGTGGTGGTAT 161
DB 2 GGAGATCAGGAGTGGAGTTGTTTCAGATCATCATCATCGTGGTGGTGGTGGTAT 61
QY 162 GGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 221
DB 62 GGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 121
QY 222 CCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 281
DB 122 CCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 181
QY 282 CTCGGAGAGCAGTGTGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 341
DB 182 CTCGGAGAGCAGTGTGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 241
QY 342 GCCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 401
DB 242 GCCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 301
QY 402 GCCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 461
DB 302 GCCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 361
QY 462 CCGGAGGAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 521
DB 362 CCGGAGGAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 421
QY 522 GCAGCTGGAAGTGAACCCGGGAGTGGTGGCGAGTGGTGGCGAGTGGTGGCGAGT 581
DB 422 GCAGCTGGAAGTGAACCCGGGAGTGGTGGCGAGTGGTGGCGAGTGGTGGCGAGT 481
QY 582 TGACCTGATGATAGTGCCAGGCTGGGCGGCCCTTGCCGCCCTTGCCGCCCTTGCC 641
DB 482 TGACCTGATGATAGTGCCAGGCTGGGCGGCCCTTGCCGCCCTTGCCGCCCTTGCC 541
QY 642 CAGCCGCACTGCTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 701
DB 542 CAGCCGCACTGCTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 601
QY 702 GGTATCGGCCACTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 761
DB 602 GGTATCGGCCACTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 661
QY 762 CTTGCTGGAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 821
DB 662 CTTGCTGGAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 721
QY 822 CTGGAGCAAGAGAGAGATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 861
DB 722 CTGGAGCAAGAGAGAGATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 761

```

QY 582 TGACCTGATGGATAGTGCAGGCTGGGCGGCCCTGCCGCCCCAGCAGTAACCTCGGCAT 641
 DB 482 TGACCTGATGGATAGTGCAGGCTGGGCGGCCCTGCCGCCCCAGCAGTAACCTCGGCAT 541
 QY 642 CAGCGCCACGTGTACTACGCGAGCGCGGCGGCGATGAGAGGCGCGCGCCACCTACAGCGA 701
 DB 542 CAGCGCCACGTGTACTACGCGAGCGCGGCGGCGATGAGAGGCGCGCGCCACCTACAGCGA 601
 QY 702 GGTTCATCGGCCCACTACCCCGGGTCTCTTCCTTCAGACACAGCAGAGAGTGGCGCCCTC 761
 DB 602 GGTTCATCGGCCCACTACCCCGGGTCTCTTCCTTCAGACACAGCAGAGAGTGGCGCCCTC 661
 QY 762 CTTGCTGAGAGGAGCCCGGCTCTACACACACATCGCGCCCTAGAGAGCGCAGCCAT 821
 DB 662 CTTGCTGAGAGGAGCCCGGCTCTACACACACATCGCGCCCTAGAGAGCGCAGCCAT 721
 QY 822 CTGGAGCAAGAGAGAGTAAACAGAAAGACACCCCTCTC 861
 DB 722 CTGGAGCAAGAGAGAGTAAACAGAAAGACACCCCTCTC 761

 RESULT 11
 LOCUS BD272516 969 bp DNA linear PAT 17-JUL-2003
 DEFINITION Secreted proteins and nucleic acids encoding them.
 ACCESSION BD272516
 VERSION BD272516.1 GI:33082284
 KEYWORDS JP 2002539773-A/25.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 969)
 AUTHORS Barnes,T.M., Holtzman,D.A., Sharp,J.D. and Fraser,C.C.
 TITLE Secreted proteins and nucleic acids encoding them
 JOURNAL Patent: JP 2002539773-A 25 26-NOV-2002;
 MILLENNium PHARMACEUTICALS INC
 COMMENT OS Homo sapiens (human)
 PN JP 2002539773-A/25
 PD 26-NOV-2002
 PF 01-MAR-2000 JP 2000502247
 PR 01-MAR-1999 US 60/122458
 PI THOMAS M BARNES,DOUGLAS A HOLTZMAN,JOHN D SHARP,CHRISTOPHER C
 PI FRASER
 PC C12N15/09,C07K14/47,C07K16/18,C12N5/10,C12P21/02,C12Q1/68, PC
 G01N33/15
 PC G01N33/50,G01N33/53,G01N33/53,G01N33/566,C12N15/00,C12N5/00 CC
 Secreted proteins and nucleic acids encoding them FH Key
 Location/Qualifiers
 FT CDS Location/Qualifiers
 (6)..(761).
 source 1..969
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

 ORIGIN
 Query Match 87.5%; Score 753.6; DB 6; Length 969;
 Best Local Similarity 99.5%; Pred. No. 7.6e-119;
 Matches 756; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 102 GGAGATCAGGAGCTGGAGTTGTTTCAGATCATCATCGTGGTGTATGTGTAT 161
 DB 2 GGAGATCAGGAGCTGGAGTTGTTTCAGATCATCATCGTGGTGTATGTGTAT 61
 QY 162 GGTGGTGGTATCAGCTGCTGTGAGCCTACAGCTCTCTGACGGTCTTCATCAG 221
 DB 62 GGTGGTGGTATCAGCTGCTGTGAGCCTACAGCTCTCTGACGGTCTTCATCAG 121
 QY 222 CCGGCACAGCCAGGGCGGAGAGAGATGGCCTGTCTCAGAGAGATGCTGTGGCC 281
 DB 122 CCGGCACAGCCAGGGCGGAGAGAGATGGCCTGTCTCAGAGAGATGCTGTGGCC 181

QY 282 CTGGAGAGCAGTGTACGCAACGGAATCCAGAGCCGAGGTCTACGCCCCGCTCG 341
 DB 182 CTGGAGAGCAGTGTACGCAACGGAATCCAGAGCCGAGGTCTACGCCCCGCTCG 241
 QY 342 GCCACCGACGGCTGGCGCTTGGCGCCCTTGGCGCCAGCGGAGCGCTTCCACGCTTCCA 401
 DB 242 GCCACCGACGGCTGGCGCTTGGCGCCCTTGGCGCCAGCGGAGCGCTTCCACGCTTCCA 301
 QY 402 GCCACCTATTCGCTACCTGCGAGCAGAGATGCACTTGCACACCCACCATCTCGTGTGAGA 461
 DB 302 GCCACCTATTCGCTACCTGCGAGCAGAGATGCACTTGCACACCCACCATCTCGTGTGAGA 361
 QY 462 CGGGAGAGAGCCGCCACCTTACAGGGCCCTTGCACCCCTTGCAGCTTGGGAGCCCGAGCA 521
 DB 362 CGGGAGAGAGCCGCCACCTTACAGGGCCCTTGCACCCCTTGCAGCTTGGGAGCCCGAGCA 421
 QY 522 GCAGCTGGAATGAAACCGGAGTGGTGGCGGACACCCCAACAGAAACCATCTTCGACAG 581
 DB 422 GCAGCTGGAATGAAACCGGAGTGGTGGCGGACACCCCAACAGAAACCATCTTCGACAG 481
 QY 582 TGACCTGATGATAGTGCAGGCTGGGCGGCCCTTGCACCCCTTGCAGCAGTAACCTCGGCGAT 641
 DB 482 TGACCTGATGATAGTGCAGGCTGGGCGGCCCTTGCACCCCTTGCAGCAGTAACCTCGGCGAT 541
 QY 642 CAGCGCCACGTGTACGCGAGCGCGGCGCATGAGAGGCGCGCCGCTACAGCGA 701
 DB 542 CAGCGCCACGTGTACGCGAGCGCGGCGCATGAGAGGCGCGCCGCTACAGCGA 601
 QY 702 GGTTCATCGGCCCACTACCCCGGGTCTCTCTTCAGACACAGCAGAGAGTGGCGCCCTC 761
 DB 602 GGTTCATCGGCCCACTACCCCGGGTCTCTCTTCAGACACAGCAGAGAGTGGCGCCCTC 661
 QY 762 CTTGCTGAGAGGAGCCCGGCTTCCACACACACATCGCGCCCTTGCAGAGCGCAGCCAT 821
 DB 662 CTTGCTGAGAGGAGCCCGGCTTCCACACACACATCGCGCCCTTGCAGAGCGCAGCCAT 721
 QY 822 CTGGAGCAAGAGAGAGTAAACAGAAAGACACCCCTCTC 861
 DB 722 CTGGAGCAAGAGAGAGTAAACAGAAAGACACCCCTCTC 761

 RESULT 12
 LOCUS AR336830 1140 bp DNA linear PAT 17-AUG-2003
 DEFINITION Sequence 1 from patent US 6566130.
 ACCESSION AR336830
 VERSION AR336830.1 GI:33722680
 KEYWORDS Unknown.
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1140)
 AUTHORS Srivastava,S., Moul,J.W., Xu,L.L. and Segawa,T.
 TITLE Androgen-regulated gene expressed in prostate tissue
 JOURNAL Patent: US 6566130-A 1 20-MAY-2003;
 FEATURES Location/Qualifiers
 source 1..1140
 /organism="unknown"
 /mol_type="genomic DNA"

 ORIGIN
 Query Match 87.5%; Score 753.6; DB 6; Length 1140;
 Best Local Similarity 99.5%; Pred. No. 7.5e-119;
 Matches 756; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 102 GGAGATCAGGAGCTGGAGTTGTTTCAGATCATCATCGTGGTGTATGTGTAT 161
 DB 91 GGAGATCAGGAGCTGGAGTTGTTTCAGATCATCATCGTGGTGTATGTGTAT 150
 QY 162 GGTGGTGGTATCAGCTGCTGTGAGCCTACAGCCTACAGCTGTCTGACGGTCTTCATCAG 221
 DB 151 GGTGGTGGTATCAGCTGCTGTGAGCCTACAGCCTACAGCTGTCTGACGGTCTTCATCAG 210

```
QY 222 CCGGACAGCCAGGCGGAGAGAGATGCCCTGTCTCTCAGAGGATGCTGTGGCC 281
Db 211 CCGGACAGCCAGGCGGAGAGAGATGCCCTGTCTCTCAGAGGATGCTGTGGCC 270
QY 282 CTCGGAGAGCAGTGTGAGCAACGGAATCCAGAGCCGAGGTCTTACGCCCGGCTCG 341
Db 271 CTCGGAGAGCAGTGTGAGCAACGGAATCCAGAGCCGAGGTCTTACGCCCGGCTCG 330
QY 342 GCCACAGAGCCGCTGGCGCTGCCCTTGGCCAGGAGGAGGCTTCCACGCTTCCA 401
Db 331 GCCACAGAGCCGCTGGCGCTGCCCTTGGCCAGGAGGAGGCTTCCACGCTTCCA 390
QY 402 GCCACCTATCGTACCTGACGACGAGATCGACTGCCACCCACCACTCTCGTGTGAGA 461
Db 391 GCCACCTATCGTACCTGACGACGAGATCGACTGCCACCCACCACTCTCGTGTGAGA 450
QY 462 CGGGAGAGAGCCGCCACCTACAGGAGCCCTGACCCCTCGAGTTCGGAGACCCGAGCA 521
Db 451 CGGGAGAGAGCCGCCACCTACAGGAGCCCTGACCCCTCGAGTTCGGAGACCCGAGCA 510
QY 522 GCAGCTGGAATGAAACCGGAGTCCGTCGGCGGACACCCGCCAAACAGAACCTATTCGACAG 581
Db 511 GCAGCTGGAATGAAACCGGAGTCCGTCGGCGGACACCCGCCAAACAGAACCTATTCGACAG 570
QY 582 TGACCTGATGATAGTGGCAGCTGGGGGGCCCTGCCCCCAGAGTAACTCGGGCAT 641
Db 571 TGACCTGATGATAGTGGCAGCTGGGGGGCCCTGCCCCCAGAGTAACTCGGGCAT 630
QY 642 CAGCGCCACGTGTACGCGACGCGGCGCATGAGAGGGCGCGGCCACCTACAGCGA 701
Db 631 CAGCGCCACGTGTACGCGACGCGGCGCATGAGAGGGCGCGGCCACCTACAGCGA 690
QY 702 GGTGATGCGGCACCTACCGGGGCTCTCTTCCAGCAGCAGAGAGTGGGGCGCCCTC 761
Db 691 GGTGATGCGGCACCTACCGGGGCTCTCTTCCAGCAGCAGAGAGTGGGGCGCCCTC 750
QY 762 CTGCTGAGAGGGAGCCCGGCTCCACCAACACACATCGCGCCCTCAGAGAGCGAGCAT 821
Db 751 CTGCTGAGAGGGAGCCCGGCTCCACCAACACACATCGCGCCCTCAGAGAGCGAGCAT 810
QY 822 CTGGAGCAAGAGAGGATTAACAGAAAGGACACCCCTCTC 861
Db 811 CTGGAGCAAGAGAGGATTAACAGAAAGGACACCCCTCTC 850

RESULT 13
AR336831 AR336831 759 bp DNA linear PAT 17-AUG-2003
LOCUS Sequence 2 from patent US 6566130.
DEFINITION AR336831
ACCESSION AR336831
VERSION AR336831.1 GI:33722681
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 759)
AUTHORS Srivastava,S., Moul,J.W., Xu,L.L. and Segawa,T.
TITLE Androgen-regulated gene expressed in prostate tissue
JOURNAL Patent: US 6566130-A 2 20-MAY-2003;
FEATURES
    source
        location/Qualifiers
            1..759
                /organism="unknown"
                /mol_type="genomic DNA"

ORIGIN
Query Match 87.4%; Score 752.8; DB 6; Length 759;
Best Local Similarity 99.7%; Pred. No. 1.1e-118;
Matches 754; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 106 ATCAGCGAGCTGGAGTTGTTTCAGATCATCATCGTGGTGGTATGATGATGATG 165
Db 1 ATGCGGAGCTGGAGTTGTTTCAGATCATCATCGTGGTGGTATGATGATGATGATG 60
```

```
QY 166 GTGGTGATCAGTGCCTGCTGAGCCACTACAGCTGTCTGACCGTCTTTCATCAGCCGG 225
Db 61 GTGGTGATCAGTGCCTGCTGAGCCACTACAGCTGTCTGACCGTCTTTCATCAGCCGG 120
QY 226 CACAGCCAGGGCGGAGGAGAGATGCCCTGTCTCAGAAAGATGCTGTGGCCCTCG 285
Db 121 CACAGCCAGGGCGGAGGAGAGATGCCCTGTCTCAGAAAGATGCTGTGGCCCTCG 180
QY 286 GAGAGCAGAGTGTAGGCAACGGATCCAGAGCGGAGGTCTAGCGCCCGCTCGGCC 345
Db 181 GAGAGCAGAGTGTAGGCAACGGATCCAGAGCGGAGGTCTAGCGCCCGCTCGGCC 240
QY 346 ACCGACCCCTGCGCGCTGCCCTTCCAGGGGAGCGCTTCCACCGCTTCCAGGCC 405
Db 241 ACCGACCCCTGCGCGCTGCCCTTCCAGGGGAGCGCTTCCACCGCTTCCAGGCC 300
QY 406 ACCTATCGTACCTGACGACGAGATCGACTGCCACCCACCACTCTGCTGACAGCGG 465
Db 301 ACCTATCGTACCTGACGACGAGATCGACTGCCACCCACCACTCTGCTGACAGCGG 360
QY 466 GAGGAGCCCGCCACCTACAGGGCCCTTGACCCCTTCCAGTTCGGGAGCCCGAGCAG 525
Db 361 GAGGAGCCCGCCACCTACAGGGCCCTTGACCCCTTCCAGTTCGGGAGCCCGAGCAG 420
QY 526 CTGGAACTGAACCGGAGTTCGGTGGCGGACACCCCAACAGAACATCTTTCAGAGTGC 585
Db 421 CTGGAACTGAACCGGAGTTCGGTGGCGGACACCCCAACAGAACATCTTTCAGAGTGC 480
QY 586 CTGATGGATAGTGCAGGCTGGGGCGCCCTGCCCCCAGCAGTAATCTCGGGCATCAGC 645
Db 481 CTGATGGATAGTGCAGGCTGGGGCGCCCTGCCCCCAGCAGTAATCTCGGGCATCAGC 540
QY 646 GCCAGTGTAGCGGAGCGGGCGCATGAGAGGGCGCGCCACCTACAGCGAGTCTC 705
Db 541 GCCAGTGTAGCGGAGCGGGCGCATGAGAGGGCGCGCCACCTACAGCGAGTCTC 600
QY 706 ATCGGCGCACTACCGGGGTCTCTTCCAGACACAGAGAGAGTGGGGCGCCCTCTCTG 765
Db 601 ATCGGCGCACTACCGGGGTCTCTTCCAGACACAGAGAGAGTGGGGCGCCCTCTCTG 660
QY 766 CTGGAGGGAGCCCGCTCCACACACATTCGCGCCCTTACAGAGCGAGCCTCTGG 825
Db 661 CTGGAGGGAGCCCGCTCCACACACATTCGCGCCCTTACAGAGCGAGCCTCTGG 720
QY 826 AGCAAGAGAGGATTAACAGAAAGGACACCCCTCTC 861
Db 721 AGCAAGAGAGGATTAACAGAAAGGACACCCCTCTC 756

AX775887 1085 bp mRNA linear PAT 14-JUL-2003
Sequence 157 from Patent WO03048202.
AX775887
AX775887.1 GI:32693605
KEYWORDS Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Matsuda,A. and Muramatsu,S.
TITLE NF-kB activating gene
JOURNAL Patent: WO 03048202-A 157 12-JUN-2003;
Asahi Kasei Kabushiki Kaisha (JP)
FEATURES
    source
        location/Qualifiers
            1..1085
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                128..886
                    /note="unnamed protein product"
                    /codon_start=1

CDS
```

```

/protein_id="CAE11643.1"
/db_xref="GI:32693606"
/translation="WAELEFVQIIIVVMMVWVITCLSHVKLARGPISHSOG
RRRDLSSBCLPSESTVSGNIPPEQVYAPRPDRLAVPPAPRERHFRQPY
PILQHEIDLPPTISDEBPYGGPCTQLRDPPEQLELNRESVAPNRITFDSD
LMSARLGGPCPPISNSGISATCYSGGRMEGPPPTYSVIGHYFGSSPQHQSSGP
SLLEGTRLHTHIAPLESAAWSKEKDKQKHPL"
ORIGIN
      87 4%; Score 752.6; DB 6; Length 1085;
      Best Local Similarity 98.2%; Pred. No. 1.1e-118;
      Matches 761; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 87 TTGTTCAGAGCATGGAGATCAGGAGCTGGAGTTTGTTCAGATCATCATCTGGT 146
DB 109 TCTCTGGAACACGAGCAATGGCGAGCTGGAGTTTGTTCAGATCATCATCTGGT 168
QY 147 GGTGATGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 206
DB 169 GGTGATGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 228
QY 207 ACGTCTCTTCATCAGCCGACAGCCAGCGGCGGAGGAGAGATGCCCTGTCTCAGA 266
DB 229 ACGTCTCTTCATCAGCCGACAGCCAGCGGCGGAGGAGAGATGCCCTGTCTCAGA 288
QY 267 AGGATGCTGTGGCTCGGAGAGACAGTGTTCAGCAACGGAATCCAGAGCCGAGGT 326
DB 289 AGGATGCTGTGGCTCGGAGAGACAGTGTTCAGCAACGGAATCCAGAGCCGAGGT 348
QY 327 CTACGCCCGCTCGGCCACAGCCGCTGGCGCTGGCGCTTCGCCCGAGCGGAGCG 386
DB 349 CTACGCCCGCTCGGCCACAGCCGCTGGCGCTGGCGCTTCGCCCGAGCGGAGCG 408
QY 387 CTTTCACCGCTTCGAGCCACCTATCCGTAACCTGAGCAGCAGATCGACCTGCCAC 446
DB 409 CTTTCACCGCTTCGAGCCACCTATCCGTAACCTGAGCAGCAGATCGACCTGCCAC 468
QY 447 CATCTCGCTGTGAGCGGAGGAGCGCCACCTACAGGCGCCCTGCACCTCCAGCT 506
DB 469 CATCTCGCTGTGAGCGGAGGAGCGCCACCTACAGGCGCCCTGCACCTCCAGCT 528
QY 507 TCGGAGCCCGAGCAGCAGCTGGAATGAAACCGGAGTGGTGGCGGACCCCGCAACAG 566
DB 529 TCGGAGCCCGAGCAGCAGCTGGAATGAAACCGGAGTGGTGGCGGACCCCGCAACAG 588
QY 567 AACCATCTTCAGAGTACCTGATAGTATGATAGTATGATAGTATGATAGTATGAT 626
DB 589 AACCATCTTCAGAGTACCTGATAGTATGATAGTATGATAGTATGATAGTATGAT 648
QY 627 CAGTAACTCGGGCATCAGCGCCACGCTGTAACGAGCGCGCGGCGCATGGAGGGCGCC 686
DB 649 CAGTAACTCGGGCATCAGCGCCACGCTGTAACGAGCGCGCGGCGCATGGAGGGCGCC 708
QY 687 GCCACCTACAGGAGTATCGGCCACTACCGGGGTCTCTTCCAGCACCAGCAGAG 746
DB 709 GCCACCTACAGGAGTATCGGCCACTACCGGGGTCTCTTCCAGCACCAGCAGAG 768
QY 747 CAGTGGCGCGCCCTCTCTTCTGGAGGGGACCGGGTCCACACACACATCGCGCCCT 806
DB 769 CAGTGGCGCGCCCTCTCTTCTGGAGGGGACCGGGTCCACACACACATCGCGCCCT 828
QY 807 AGAGAGCGAGCATCTGAGCGAAGAGAGATGATGATGATGATGATGATGATGATGAT 861
DB 829 AGAGAGCGAGCATCTGAGCGAAGAGAGATGATGATGATGATGATGATGATGATGAT 883

```

```

RESULT 15
BD272545 LOCUS 759 bp DNA linear PAT 17-JUL-2003
DEFINITION Secreted proteins and nucleic acids encoding them.
ACCESSION BD272545
VERSION BD272545.1 GI:33082313
KEYWORDS JP 2002539773-A/54.

```

```

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 759)
AUTHORS Barnes, T.M., Holtzman, D.A., Sharp, J.D. and Fraser, C.C.
TITLE Secreted proteins and nucleic acids encoding them
JOURNAL Patent: JP 2002539773-A 54 26-NOV-2002;
MILLENNIUM PHARMACEUTICALS INC
COMMENT OS Homo sapiens (human)
PN JP 2002539773-A/54
PD 26-NOV-2002 JP 2000602247
PF 01-MAR-2000 JP 60/122458
PI THOMAS M BARNES, DOUGLAS A HOLTZMAN, JOHN D SHARP, CHRISTOPHER C
PI FRASER
PC C12N15/09, C07K14/47, C07K16/18, C12N5/10, C12P21/02, C12Q1/68, PC
G01N33/15,
PC G01N33/50, G01N33/53, G01N33/53, G01N33/56, C12N15/00, C12N5/00 CC
Secreted proteins and nucleic acids encoding them FH Key
Location/Qualifiers
FT source 1..759
FT /organism="Homo sapiens"
FT /mol_type="genomic DNA"
FT /db_xref="taxon:9606"
FEATURES
source
Query Match 87.4%; Score 752.4; DB 6; Length 759;
Best Local Similarity 99.6%; Pred. No. 1.3e-118;
Matches 753; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
ORIGIN
      106 ATCAAGAGCTGGAGTTTGTTCAGATCATCATCTGGTGGTGGTGGTGGTGGTGGT 165
      1 ATGCGGAGCTGGAGTTTGTTCAGATCATCATCTGGTGGTGGTGGTGGTGGTGGT 60
      166 GTGGTATCATCTGCTGAGCCACTACAAGCTGTCTGACGGTCTTCATCAGCCGG 225
      61 GTGGTATCATCTGCTGAGCCACTACAAGCTGTCTGACGGTCTTCATCAGCCGG 120
      226 CACAGCCAGGGCGGAGGAGAGATGCTCTCTCAGAAAGATGCTGTGGGCTTCG 285
      121 CACAGCCAGGGCGGAGGAGAGATGCTCTCTCAGAAAGATGCTGTGGGCTTCG 180
      286 GAGAGCAGCTGTCAGCAACGGAATCCAGAGCCGAGCTACGCCCGCTTCGCCCC 345
      181 GAGAGCAGCTGTCAGCAACGGAATCCAGAGCCGAGCTACGCCCGCTTCGCCCC 240
      346 ACCGACCGCTGGCGCTGGCGCTTCGCCCGAGCGGAGCGCTTCCACCGCTTCAG 405
      241 ACCGACCGCTGGCGCTGGCGCTTCGCCCGAGCGGAGCGCTTCCACCGCTTCAG 300
      406 ACCTATCGTACCTGACGACGAGATCGACCTGCGACCCACCATCTCGCTGTGAG 465
      301 ACCTATCGTACCTGACGACGAGATCGACCTGCGACCCACCATCTCGCTGTGAG 360
      466 GAGAGCGCCCGACCTACCAAGGCGCTGCACTCTCAGCTTCGGGACCCCGAGCAG 525
      361 GAGAGCGCCCGACCTACCAAGGCGCTGCACTCTCAGCTTCGGGACCCCGAGCAG 420
      526 CTGGAATCTGAACCGGAGTGGTGGCGGACCCCGGACCCCGGACCATCTTCGAG 585
      421 CTGGAATCTGAACCGGAGTGGTGGCGGACCCCGGACCCCGGACCATCTTCGAG 480
      586 CTGATGATAGTGCAGGCTGGGCGGCGCTGCGCCCGGACCATCTTCGAGCATCAG 645
      481 CTGATGATAGTGCAGGCTGGGCGGCGCTGCGCCCGGACCATCTTCGAGCATCAG 540
      646 GCCACGTGCTACGCGACGCGCGCGCATGAGAGGGCGCGCGCCACCTACAGCGAG 705
      541 GCCACGTGCTACGCGACGCGCGCGCATGAGAGGGCGCGCGCCACCTACAGCGAG 600

```


GenCore version 5.1.6
Copyright (C) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 11, 2004, 21:06:02 ; Search time 599,354 Seconds
(without alignments)
9363.200 Million cell updates/sec

Title: US-09-934-249-1

Perfect score: 1321

Sequence: 1 cgacgcggtctcgagcga.....ctgcgtaggtgaaagggcag 1321

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002s.*
7: Geneseqn2003as.*
8: Geneseqn2003bs.*
9: Geneseqn2003cs.*
10: Geneseqn2004s.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1321	100.0	1321	6	ABK12137 Human CDN
2	1229	93.0	4839	7	ACC49552 Tumour-as
3	1227.4	92.9	1383	9	ADC37326 Nuclear f
4	803.2	60.8	969	3	AAA75151 CDNA enco
5	802.2	60.7	1140	6	ABK92120 Prostate
6	802.2	60.7	1141	9	ADC75588 Prostate
7	802.2	60.7	1850	7	ACC49536 Tumour-as
8	801.6	60.7	969	3	AAA75164 CDNA clon
9	801.6	60.7	969	3	AAA75163 CDNA clon
10	801.6	60.7	969	3	AAA75185 CDNA clon
11	801.6	60.7	1140	9	AAD60105 Human and
12	800.6	60.6	1061	3	AAA47429 Sequence
13	800.6	60.6	1085	9	ADC37324 Nuclear f
14	800.6	60.6	1334	7	ABZ36103 Human sec
15	790.4	59.8	1066	4	AAI57868 Human pol
16	767.4	58.1	806	7	ACC49537 Tumour-as
17	760	57.5	1089	4	AAI59654 Human pol
18	616.4	46.7	878	6	ABK12142 Mouse CDN
19	589.2	44.6	1583	6	ABSG1424 Prostate
20	493.6	37.4	693	6	ABK12143 Human MIV
21	466.2	35.3	1713	3	AAA75152 CDNA enco
22	464.6	35.2	1713	3	AAA75167 CDNA clon
23	464.6	35.2	1713	3	AAA75166 CDNA clon

24	454.6	35.2	1713	3	AAA75168	AAA75168 CDNA clon
25	417.8	31.6	812	2	AAZ52964	AAZ52964 Human pro
26	417	31.6	474	7	ABZ84732	ABZ84732 Toxicolog
c 27	401.4	30.4	408	5	AAF65983	AAF65983 Novel hum
28	400.6	30.3	673	6	ABT09178	ABT09178 Phase-1 R
29	352.2	26.7	8494	5	AAZ77304	AAZ77304 DNA enco
30	351.6	26.6	937	3	AAZ52507	AAZ52507 Human sec
31	350.4	26.5	1879	5	AAZ84503	AAZ84503 DNA enco
32	350	26.5	920	6	ABK34251	ABK34251 Human CDN
33	350	26.5	2170	7	ADA52629	ADA52629 Human cod
34	349	26.4	8065	2	AAV38335	AAV38335 Manic-dep
35	349	26.4	8093	6	ABK12145	ABK12145 Human MIV
36	341.8	25.9	475	6	ABK12144	ABK12144 Human MIV
c 37	333.8	25.3	8103	5	AAZ77312	AAZ77312 DNA enco
c 38	298	22.6	837	6	ABQ43498	ABQ43498 Oligonuc
39	298	22.5	837	6	ABQ43499	ABQ43499 Oligonuc
40	284.2	21.5	1496	9	ADC37452	ADC37452 Nuclear f
41	284.2	21.5	8440	5	AAZ77305	AAZ77305 DNA enco
42	284.2	21.5	8440	6	ABK83477	ABK83477 Human CDN
43	281	21.3	8011	2	AAV38336	AAV38336 Manic-dep
c 44	270.8	20.5	837	6	ABQ43501	ABQ43501 Oligonuc
45	270.8	20.5	837	6	ABQ43500	ABQ43500 Oligonuc

ALIGNMENTS

RESULT 1
ABK12137
ID ABK12137 standard; CDNA; 1321 BP.

AC ABK12137;

DT 05-JUN-2002 (first entry)

XX Human cDNA encoding Mechanically Induced Vascular Receptor 1, MIVR-1.

XX Human; ss; gene; MIVR-1; Mechanically Induced Vascular Receptor 1;

KW cytosolic; cardiant; cerebroprotective; antiarteriosclerotic;

KW cardiac cell; anti-apoptotic; vascular endothelial cell;

KW cardiac hypertrophy; myocardial infarction; stroke; arteriosclerosis;

XX heart failure.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 413..1276

FT /*tag= a

FT /product= "MIVR-1"

FT /note= "This region is specifically claimed in claim 3"

XX WO200216416-A2.

XX 28-FEB-2002.

XX 21-AUG-2001; 2001WO-US026089.

XX 22-AUG-2000; 2000US-0227159P.

XX (BGHM) BRIGHAM & WOMENS HOSPITAL INC.

XX (PFIZ) PFIZER INC.

XX Lee RT, Landschulz KT, Kennedy SP, Thompson JF, Turi TG;

XX WPI, 2002-280912/32.

XX P-P8DB; AAU78231.

XX Novel nucleic acid molecule encoding Mechanically Induced Vascular

XX Receptor-1 polypeptide, useful for treating cardiovascular diseases.

XX Claim 2; Page 87-88; 105pp; English.

XX The invention relates to an isolated nucleic acid molecule encoding a

CC Mechanically Induced Vascular Receptor (MIVR)-1 polypeptide having
CC cardiac cell anti-apoptotic activity and fragments of it provided they
CC are not identical to Genbank sequences A1761441.1, A1594390, NM 004338
CC and A0177451. Also included are expression vectors, host cells, the MIVR-
CC 1 polypeptide, MIVR-1 binding peptides, modulators of MIVR-1, contacting
CC a molecule having cardiac cell anti-apoptotic activity with a candidate
CC agent, where the molecule is a nucleic acid molecule comprising MIVR-1,
CC IEX-1, VDU-1, BTG-2 and TIS-11d or its expression product, determining
CC if the anti-apoptotic activity is modulated and thereby identifying a
CC modulator. The cardiac cell anti-apoptotic molecules and nucleic acids
CC of the invention are useful for treating, diagnosing and monitoring
CC progression of such diseases and disorders as characterized by increased
CC apoptotic cell-death of vascular endothelial cells e.g. cardiac
CC hypertrophy, myocardial infarction, stroke, arteriosclerosis and heart
CC failure. The present sequence encodes human MIVR-1
XX
SQ Sequence 1321 BP; 223 A; 493 C; 440 G; 165 T; 0 U; 0 Other;

Query Match 100.0%; Score 1321; DB 6; Length 1321;
Best Local Similarity 100.0%; Pred. No. 1.9e-198;
Matches 1321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACCGCGGTCTCGGAGGAAACCCGATCTCTTGGACTTGAATGAGGAGGAGGCGG 60
DB 1 CGACCGCGGTCTCGGAGGAAACCCGATCTCTTGGACTTGAATGAGGAGGAGGCGG 60
QY 61 CGGCGCGCGCGCGCGGAGGCGCTCGGCTGGGAAAGCTAGCGCAGAGCTAGCCCC 120
DB 61 CGGCGCGCGCGCGCGGAGGCGCTCGGCTGGGAAAGCTAGCGCAGAGCTAGCCCC 120
QY 121 GCGCGCGAGCGCGCGCGCGCTCGGAGGCGGCGGCGGCGGCGGCGGCGGCGG 180
DB 121 GCGCGCGAGCGCGCGCGCGCTCGGAGGCGGCGGCGGCGGCGGCGGCGGCGG 180
QY 181 AGCGCGCGCGCGCGCGGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
DB 181 AGCGCGCGCGCGCGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
QY 241 TGAGCGCGCGCGCGCGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
DB 241 TGAGCGCGCGCGCGCGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
QY 301 CTCTCCCGCGCGCGCGCTCTCGATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
DB 301 CTCTCCCGCGCGCGCGCTCTCGATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
QY 361 CGCGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
DB 361 CGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
QY 421 CTGTATGGGGGTCAACAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
DB 421 CTGTATGGGGGTCAACAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
QY 481 GTGCACTCAACAGCGCTCTTGTTCAGAGCATGGAGATCAGGAGCTGGAGTTGTTC 540
DB 481 GTGCACTCAACAGCGCTCTTGTTCAGAGCATGGAGATCAGGAGCTGGAGTTGTTC 540
QY 541 GATCATCATCATCGT 600
DB 541 GATCATCATCATCGT 600
QY 601 CCACTACAGCTGTCTGACCGTCTTTCATCAGCGGACAGCGGCGGCGGCGGAGAGA 660
DB 601 CCACTACAGCTGTCTGACCGTCTTTCATCAGCGGACAGCGGCGGCGGAGAGA 660
QY 661 AGATGCGCTGTCTCAGAGGAGTGTCTGTGGCGCGCGCGCGCGCGCGCGCGCGCG 720
DB 661 AGATGCGCTGTCTCAGAGGAGTGTCTGTGGCGCGCGCGCGCGCGCGCGCGCGCG 720
QY 721 AATCCAGAGCGCGAGGTCTACCGCGCGCGTCTGGCGCGCGCGCGCGCGCGCGCGCG 780
DB 721 AATCCAGAGCGCGAGGTCTACCGCGCGCGTCTGGCGCGCGCGCGCGCGCGCGCGCG 780

QY 781 CTTGCGCCAGCGGAGCGGCTTCCACCGCTTCCAGGCCACCTATCCGTACCTCAGACGA 840
DB 781 CTTGCGCCAGCGGAGCGGCTTCCACCGCTTCCAGGCCACCTATCCGTACCTCAGACGA 840
QY 841 GATCGACCTGCGACCCACCATCTCTGCTGTCAGACGGGAGGAGCCGCCACCTACAGGG 900
DB 841 GATCGACCTGCGACCCACCATCTCTGCTGTCAGACGGGAGGAGCCGCCACCTACAGGG 900
QY 901 CCGCTGACCCCTCCAGCTTTCGGGAGCCCGAGCAGCAGCTTGAACCCGGAGTCGT 960
DB 901 CCGCTGACCCCTCCAGCTTTCGGGAGCCCGAGCAGCAGCTTGAACCCGGAGTCGT 960
QY 961 GCGCGACCCCGGAAACAGAACCATCTTCCAGACGTACCTGATGGATGTCAGGCTGG 1020
DB 961 GCGCGACCCCGGAAACAGAACCATCTTCCAGACGTACCTGATGGATGTCAGGCTGG 1020
QY 1021 CGGCGCGCTGCGCGCGCGCGGAGTAACCTCGGCGCATCAGCGCCACCTGCTACGGCAGCGGG 1080
DB 1021 CGGCGCGCTGCGCGCGCGCGGAGTAACCTCGGCGCATCAGCGCCACCTGCTACGGCAGCGGG 1080
QY 1081 GCGCATGGAGGGGCGCGCGCGCGCGCATCTAGACGAGTCTATCGGCCACTACCGGGTCCCTC 1140
DB 1081 GCGCATGGAGGGGCGCGCGCGCGCGCATCTAGACGAGTCTATCGGCCACTACCGGGTCCCTC 1140
QY 1141 CTTCCAGCAGCAGCAGCAGCAGTGGGCGCGCGCTCTTCTGCTGGAGGGGACCCCGGCTCCACCA 1200
DB 1141 CTTCCAGCAGCAGCAGCAGCAGTGGGCGCGCGCTCTTCTGCTGGAGGGGACCCCGGCTCCACCA 1200
QY 1201 CACACATCGCGCGCGCGCTAGAGAGCGCAGCATCTGGAGCAAGAGAGAGATTAACAGAA 1260
DB 1201 CACACATCGCGCGCGCGCTAGAGAGCGCAGCATCTGGAGCAAGAGAGAGATTAACAGAA 1260
QY 1261 AGGACACCTCTCTAGGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1320
DB 1261 AGGACACCTCTCTAGGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1320
QY 1321 G 1321
DB 1321 G 1321

RESULT 2
ACC49552
ID ACC49552 standard; cDNA; 4839 BP.
XX ACC49552;
XX AC ACC49552;
XX DF 01-JUL-2003 (first entry)
XX DE Tumour-associated antigenic target protein TAT379 SEQ ID NO:119.
XX KW Human; tumour-associated antigenic target; TAT; tumour; diagnosis;
XX KW cancer; gene; ss.
XX OS Homo sapiens.
XX XX
XX FN WO2003024392-A2.
XX XX
XX PD 27-MAR-2003.
XX XX
XX PF 11-SEP-2002; 2002WO-US028859.
XX XX
XX PR 18-SEP-2001; 2001US-0323268P.
XX PR 19-OCT-2001; 2001US-0339227P.
XX PR 07-NOV-2001; 2001US-0336827P.
XX PR 20-NOV-2001; 2001US-0331906P.
XX PR 02-JAN-2002; 2002US-0345444P.
XX PR 03-APR-2002; 2002US-0369724P.
XX PR 19-AUG-2002; 2002US-0404809P.
XX (GETH) GENENTECH INC.
XX PA
XX XX

PI	Prantz G, Hillan KJ, Phillips HS, Polakis P, Spencer SD;
PI	Williams EW, Wu TD, Zhang Z;
XX	WPI; 2003-354551/33.
DR	P-PSDB; ABP97234.
XX	
PT	New antibodies against tumor-associated antigenic target polypeptide,
PT	useful for treating or diagnosing tumors or cancers in mammals, e.g.
PT	prostate cancer, lung cancer, prostate adenocarcinomas or renal cell
PT	carcinomas.
XX	
PS	Claim 2; Fig 119; 285pp; English.
XX	
CC	ACC49493 to ACC49552 encode the human tumour-associated antigenic target
CC	(TAT) proteins given in ABP97175 to ABP97234. The present invention
CC	describes an isolated antibody that binds to a polypeptide having at
CC	least 80 % sequence identity to any of the 60 150-500 residue amino acid
CC	sequences (S1), given in ABP97175 to ABP97234, comprising (S1), lacking
CC	its associated signal peptide, encoded by any of the 60 2000-3000 base
CC	pair sequences (S2), given in ACC49493 to ACC49552. TAT proteins have
CC	cytostatic activity. The antibody can be used for treating or diagnosing
CC	tumours or cancers in mammals, e.g. prostate cancer, lung cancer, breast
CC	cancer, colon cancer, ovarian cancer, prostate adenocarcinomas, renal
CC	cell carcinomas, or thyroid cancer
XX	
XX	Sequence 4839 BP; 1103 A; 1313 C; 1290 G; 1133 T; 0 U; 0 Other;
XX	
Qy	Query Match 93.0%; Score 1229; DB 7; Length 4839;
Db	Best Local Similarity 100.0%; Pred. No. 4.6e-184;
Qy	Matches 1229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	93 GGAAAGCTAGCGGCGAGGCTCAGCCCGCGGGAGCGCGCCCGCTGCCAGGCCCATTT 152
Db	1 GGAAAGCTAGCGGCGAGGCTCAGCCCGCGGGAGCGCGCCCGCTGCCAGGCCCATTT 60
Qy	153 TTCCGGAGCGCACCCCGCGGCACCTGCCAGACCCCGCGGGGCTGCCAGGAGGAGCGCGGG 212
Db	61 TTCCGGAGCGCACCCCGCGGCACCTGCCAGACCCCGCGGGGCTGCCAGGAGGAGCGCGGG 120
Qy	213 GGGCGAGCGAGCGCGTTCCTCGGCGACTGAGCCCCCGCGCGCCCGCGGAACTTGGCGGC 272
Db	121 GGGCGAGCGAGCGCGTTCCTCGGCGACTGAGCCCCCGCGCGCCCGCGGAACTTGGCGGC 180
Qy	273 GACCCGAGCCGGCGAGCGGGGCGCGCTCCCGCGCGCGCCCTTCGCAATGCGGGCG 332
Db	181 GACCCGAGCCGGCGAGCGGGGCGCGCTCCCGCGCGCGCCCTTCGCAATGCGGGCG 240
Qy	333 CCCAGTCCGGCGCGCCCGCGAGCCCCCGCGCGCCCGCGAGCCCCCGCGCCCGCGC 392
Db	241 CCCAGTCCGGCGCGCCCGCGAGCCCCCGCGCGCCCGCGAGCCCCCGCGCCCGCGC 300
Qy	393 GCGCGCGCGCGCGCCGCTCATGACCGCTTGATGGGGGTCAACAGCACCGCGCGCGCG 452
Db	301 GCGCGCGCGCGCGCCGCTCATGACCGCTTGATGGGGGTCAACAGCACCGCGCGCGCG 360
Qy	453 CGCGCGGCGAGCCCAATGTCTCTTGACGTGCAACTGCAAACGCTCTTTGTTCCAGAGCA 512
Db	361 CGCGCGGCGAGCCCAATGTCTCTTGACGTGCAACTGCAAACGCTCTTTGTTCCAGAGCA 420
Qy	513 TGGAGATACGAGCTGGAGTTTGTTCAGATCATCATCTCGTGGTGGTGGATGGTGA 572
Db	421 TGGAGATACGAGCTGGAGTTTGTTCAGATCATCATCTCGTGGTGGTGGATGGTGA 480
Qy	573 TGTGGTGGTATCATCGTGCCTGTGAGCCACTCAAGCTGTCTGCACGGTCTCTTCATCA 632
Db	481 TGTGGTGGTATCATCGTGCCTGTGAGCCACTCAAGCTGTCTGCACGGTCTCTTCATCA 540
Qy	633 GCGGCGACACCGAGGGCGGAGAGAGATGCCCTGTCTCTCAGAGAGATGCTGTGGC 692
Db	541 GCGGCGACACCGAGGGCGGAGAGAGATGCCCTGTCTCTCAGAGAGATGCTGTGGC 600
Qy	693 CTCGAGAGACACAGTGTTCAGGCAACCGAAATCCAGAGCGCGAGGCTTCAGCCCCGCTC 752

601	CTCGGAGGACACAGTGTACGACACGAAATCCACAGACCGCAGCTCTACGCCCGCCCTC	660
753	GGCCACACGACCGCTGCGCGGTGCCCGCTTTCGCCACGAGGAGCGCTTTCACACCGTTCC	812
661	GGCCCCACGACCGCTGCGCGGTGCCCGCTTTCGCCACGAGGAGCGCTTTCACACCGTTCC	720
813	AGCCCCACCTATCCGTACCTGACGACGAGATCGACTCTGCCACCCACCATCTCGCTGTCA	872
721	AGCCCCACCTATCCGTACCTGACGACGAGATCGACTCTGCCACCCACCATCTCGCTGTCA	780
873	ACGGGGAGAGGCCCCACCTTACAGAGGCGCTTGCACCTTCAGCTTCGGGACCCCGAGC	932
781	ACGGGGAGAGGCCCCACCTTACAGAGGCGCTTGCACCTTCAGCTTCGGGACCCCGAGC	840
933	AGCAGCTGGAATGAACCGGGAGTCTGCTGCGCGCACCCGCCAACAGAACCATCTTCGACA	992
841	AGCAGCTGGAATGAACCGGGAGTCTGCTGCGCGCACCCGCCAACAGAACCATCTTCGACA	900
993	GTGACCTGATGATAGTATGCTGAGGCTGGGGGCGCTTGCACCTTCAGCTTCGGGCA	1052
901	GTGACCTGATGATAGTATGCTGAGGCTGGGGGCGCTTGCACCTTCAGCTTCGGGCA	960
1053	TCAGCGCCACAGTGTCTACGACGCGCGGGCGCATGGAGGGGCGCGCCACCTTCAGCG	1112
961	TCAGCGCCACAGTGTCTACGACGCGCGGGCGCATGGAGGGGCGCGCCACCTTCAGCG	1020
1113	AGGTATCGGCCACTACCGGGGTCTCTTCACGACACGACGAGCAGTGGCGCGCCT	1172
1021	AGGTATCGGCCACTACCGGGGTCTCTTCACGACACGACGAGCAGTGGCGCGCCT	1080
1173	CTTGTCTGAGGGGACCCCGCTTCCACACACACATCGCGCCCTTAGAGAGCGCAGCCA	1232
1081	CTTGTCTGAGGGGACCCCGCTTCCACACACACATCGCGCCCTTAGAGAGCGCAGCCA	1140
1233	TTTGGAGCAAGAGAGGATTAACAGAAAGGACACCTCTCTAGGGTCCCGAGGGGGCC	1292
1141	TTTGGAGCAAGAGAGGATTAACAGAAAGGACACCTCTCTAGGGTCCCGAGGGGGCC	1200
1293	GGGCTGGGCTCGTAGGTGAAAGGCAG	1321
1201	GGGCTGGGCTCGTAGGTGAAAGGCAG	1229
RESULT 3		
ADC37326		
ID ADC37326 standard; DNA; 1383 BP.		
XX	ADC37326;	
XX	AC	
XX	AC	
XX	DT	
XX	18-DEC-2003 (first entry)	
XX	Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 159.	
XX	Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;	
XX	cancer; infectious disease; bone disease; AIDS;	
KW	neurodegenerative disease; ischemic disorder; Antinflammatory;	
KW	Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-Hiv;	
KW	Neuroprotective; Nootropic; Cardiant; Gene therapy; human; gene; ds.	
XX		
XX	Homo sapiens.	
XX		
XX	WO2003048202-A2.	
XX		
XX	12-JUN-2003.	
XX		
XX	03-DEC-2002; 2002WO-JP012644.	
XX		
XX	03-DEC-2001; 2001JP-00368692.	
PR	05-DEC-2001; 2001US-035829P.	
PR	03-OCT-2002; 2002JP-00291302.	
PR	04-OCT-2002; 2002US-0415769P.	
XX		
XX	(ASAH) ASAH KASEI KK.	

XX Matsuda A, Muramatsu S;
 PI WPI: 2003-505282/47.
 DR P-PSDB; ADC37327.
 XX
 PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
 PT useful for treating inflammation, autoimmune diseases, cancers,
 PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
 PT ischaemic disorders.
 XX
 XX Claim 4; SEQ ID NO 159; 938pp; English.
 XX
 CC The present invention relates to novel proteins and their coding
 CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
 CC kappaB). The proteins and their coding sequences are useful for treating
 CC a disease associated with NF-kappaB activation, such as inflammation,
 CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
 CC neurodegenerative diseases, or ischaemic disorders.
 XX
 SQ Sequence 1383 BP; 253 A; 505 C; 440 G; 185 T; 0 U; 0 Other;
 Query Match 92.9%; Score 1227.4; DB 9; Length 1383;
 Best Local Similarity 99.9%; Pred. No. 8.8e-184;
 Matches 1228; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 93 GGAAGCTAGCGGAGAGGCTCAGCCCGCGGCGGAGCGCGCCCGCTGCCAGCCATT 152
 Db 1 GGAAGCTAGCGGAGAGGCTCAGCCCGCGGCGGAGCGCGCCCGCTGCCAGCCATT 60
 153 TTCCGAGCCACCCCGCGGCGACTGCCAGACCCCGCGGCGTCCGAGGCGAGCGCGGG 212
 Db 61 TTCCGAGCCACCCCGCGGCGACTGCCAGACCCCGCGGCGTCCGAGGCGAGCGCGGG 120
 213 GGGCGCAGCGAGCGGTCCTCCGCGCACTGAGCCCGCGGCGCCCGCGGCACTGGCGGC 272
 Db 121 GGGCGCAGCGAGCGGTCCTCCGCGCACTGAGCCCGCGGCGCCCGCGGCACTGGCGGC 180
 273 GACCCGAGCCCGCGGCGGCGGCTCCCGCGCGCGCGCTCTCTGATCGCGGGC 332
 Db 181 GACCCGAGCCCGCGGCGGCGGCGCTCCCGCGCGCGCGCTCTCTGATCGCGGGC 240
 333 CCCAGCTCCGGCGCGCGGAGCGCCCGCGGCGCGCGCGCGCGCGCGCGCGCGC 392
 Db 241 CCCAGCTCCGGCGCGCGGAGCGCCCGCGGCGCGCGCGCGCGCGCGCGCGCGC 300
 393 GCGCGCGCGCGCGCGCTCCATGACCGCTTGATGGGGGTCAACAGACCGCGCGCGCG 452
 Db 301 GCGCGCGCGCGCGCGCTCCATGACCGCTTGATGGGGGTCAACAGACCGCGCGCG 360
 453 CCGCGCGGAGCCCAATGTCTCTGCGCACTGCAAGTCAAAAGCTTTTGTTCAGAGCA 512
 Db 361 CCGCGCGGAGCCCAATGTCTCTGCGCACTGCAAGTCAAAAGCTTTTGTTCAGAGCA 420
 513 TGGAGATCAGCGAGCTGAGTTTGTTCAGATCATCATCTGCTGCTGATGATGTTGA 572
 Db 421 TGGAGATCAGCGAGCTGAGTTTGTTCAGATCATCATCTGCTGCTGATGATGTTGA 480
 573 TGGTGGTGTATCATCGTCCCTGCTGAGCCACTACAAGCTGTCTGACGCTCTTATCA 632
 Db 481 TGGTGGTGTATCATCGTCCCTGCTGAGCCACTACAAGCTGTCTGACGCTCTTATCA 540
 633 GCGCGCACAGCGCGGCGGAGAGAGATGCCCTGCTCTCAGAGGATGCTGTGGC 692
 Db 541 GCGCGCACAGCGCGGCGGAGAGAGATGCCCTGCTCTCAGAGGATGCTGTGGC 600
 693 CTTCCGAGAGCAGTGTTCAGGCAACCGGAATCCAGAGCCGCGAGCTTACGCCCGCGCTC 752
 Db 601 CTTCCGAGAGCAGTGTTCAGGCAACCGGAATCCAGAGCCGCGAGCTTACGCCCGCGCTC 660
 753 GGGCCACCCAGCGCTGGCGCTGCGCGCTTCCCGCAGGAGCGCTTCCACCGCTTCC 812
 Db 661 GGGCCACCCAGCGCTGGCGCTGCGCGCTTCCCGCAGGAGCGCTTCCACCGCTTCC 720

QY 813 AGCCACCTATCCGTACCTGACGACGAGATCGACCTGCCACCCACCATCTCGTGTGAG 872
 Db 721 AGCCACCTATCCGTACCTGACGACGAGATCGACCTGCCACCCACCATCTCGTGTGAG 780
 QY 873 AGGGGAGAGCCCGCCACCTTACGAGGCGCCCTGACCCCTCCAGCTTCGGGACCCCGGAGC 932
 Db 781 AGGGGAGAGCCCGCCACCTTACGAGGCGCCCTGACCCCTCCAGCTTCGGGACCCCGGAGC 840
 QY 933 AGCAGCTGAACTGAAACCGGAGTGGGTGCGCGACACCCCGCCAAACAGAACCATCTTCGACA 992
 Db 841 AGCAGCTGAACTGAAACCGGAGTGGGTGCGCGACACCCCGCCAAACAGAACCATCTTCGACA 900
 QY 993 GTGACCTGATGATAGTGCAGGCTGGGCGGCGCCCTGCCCGCCAGCAGTAACCTCGGCA 1052
 Db 901 GTGACCTGATGATAGTGCAGGCTGGGCGGCGCCCTGCCCGCCAGCAGTAACCTCGGCA 960
 QY 1053 TCAGCGCCACCTGTCTTACCGGAGCGCGCGCGCATGAGGCGCGCCCGCCACCTTACAGCG 1112
 Db 961 TCAGCGCCACCTGTCTTACCGGAGCGCGCGCGCATGAGGCGCGCCCGCCACCTTACAGCG 1020
 QY 1113 AGGTATCGGCCACTTACCGGCGGTCTCTCTCCAGCACCAGCAGAGTGGGCGCGCCT 1172
 Db 1021 AGGTATCGGCCACTTACCGGCGGTCTCTCTCCAGCACCAGCAGAGTGGGCGCGCCT 1080
 QY 1173 CTTTCTGAGGGGACCCCGGCTCCACACACACACATCGCGCCCTTACAGAGCGCAGCA 1232
 Db 1081 CTTTCTGAGGGGACCCCGGCTCCACACACACACATCGCGCCCTTACAGAGCGCAGCA 1140
 QY 1233 TCTGAGCAAGAGAGGATTAACAGAAAGACACCTCTCTAGGCTCCACAGGGGGCC 1292
 Db 1141 TCTGAGCAAGAGAGGATTAACAGAAAGACACCTCTCTAGGCTCCACAGGGGGCC 1200
 QY 1293 GGGCTGGGCTCGGTAGGTGAAAGGCGAG 1321
 Db 1201 GGGCTGGGCTCGGTAGGTGAAAGGCGAG 1229
 RESULT 4
 AAA75151
 ID AAA75151 standard; cDNA; 969 BP.
 XX
 AC AAA75151;
 XX
 DT 15-JAN-2001 (first entry)
 XX
 DE cDNA encoding a human TANGO 261 polypeptide.
 XX
 KW TANGO 266; TANGO 261; TANGO 262; TANGO 267;
 KW cellular proliferation; cellular differentiation; cellular adhesion;
 KW von Willebrand factor-associated disorder; cell trafficking; cancer;
 KW hematopoietic associated disease; atelectasis; pulmonary congestion;
 KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;
 KW intestinal disorder; spleen associated disease; renal disorder;
 KW cardiovascular disorder; ischemic heart disease; hydrocephalus;
 KW brain herniation; iatrogenic disease; inflammation; meningitis;
 KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;
 KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 CD 6..764
 FT /*tag= a
 FT /product= "TANGO 261"
 FT 6..89
 FT /*tag= b
 FT mat_peptide 90..764
 FT /*tag= c
 XX
 PN WO200052022-A1.
 XX
 PD 08-SEP-2000.

PT prostate tissue.
 PS Claim 22; Page 305; 436pp; English.
 XX The present invention relates to methods of detecting a prostate cancer-associated transcript in a cell from a patient. The method comprises contacting a biological sample from the patient with prostate cancer-associated polynucleotides (designated PC genes) that selectively hybridise to a sequence that is at least 80% identical to them. The prostate cancer-associated polynucleotide sequences are differentially expressed in prostate tumour tissue or in prostate cancer and are derived from the tissues of various organisms such as humans or other mammals (e.g. mice, sheep and dogs). The methods of the invention are useful for diagnosing and treating prostate cancer in mammals. The prostate cancer-associated genes are useful for diagnosing or treating prostate cancer, as well as for identifying modulators of prostate cancer or agents that inhibit prostate cancer. The nucleic acid sequences are particularly useful in gene therapy, as a vaccine or in antisense applications.
 CC ABK92115-ABK92263 represent prostate cancer-associated polynucleotide sequences
 XX
 SQ Sequence 1140 BP; 270 A; 350 C; 336 G; 184 T; 0 U; 0 Other;
 Query Match 60.7%; Score 802.2; DB 6; Length 1140;
 Best Local Similarity 98.4%; Pred. No. 3.9e-117;
 Matches 810; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
 QY 499 TTGTTCAGAGCTGGAGATCAGGAGCTGGAGTTTGTTCAGATCATCATCGTGGT 558
 DB 77 TCTCCTGCAAAACAGGCAATGGGGAGCTGGAGTTTGTTCAGATCATCATCGTGGT 136
 QY 559 GGTGATGATGTGATGGTGGTGTATCATCAGTGTCTGTGAGCCACTACAAAGCTGTCTGC 618
 DB 137 GGTGATGATGTGATGGTGGTGTATCATCAGTGTCTGTGAGCCACTACAAAGCTGTCTGC 196
 QY 619 ACGGTCTTTCATCAGCCGGCACAGCCAGGGCGGAGGAGAGAGATGCCCTGTCTCTCAGA 678
 DB 197 ACGGTCTTTCATCAGCCGGCACAGCCAGGGCGGAGGAGAGAGATGCCCTGTCTCTCAGA 256
 QY 679 AGGATGCTGTGGCTCGGAGAGCACAGTGTCTGAGCAACGGATCCAGAGCCGAGGT 738
 DB 257 AGGATGCTGTGGCTCGGAGAGCACAGTGTCTGAGCAACGGATCCAGAGCCGAGGT 316
 QY 739 CTAGCCCGGCTCGGCCACCGAGCCGCTGCGCGTGGCGCCCTTCGCCAGCGGGAGCG 798
 DB 317 CTAGCCCGGCTCGGCCACCGAGCCGCTGCGCGTGGCGCCCTTCGCCAGCGGGAGCG 376
 QY 799 CTTCCACCGCTTCAGCCCACTATCTGCTGACGACAGATCGACCTGCCACCCAC 858
 DB 377 CTTCCACCGCTTCAGCCCACTATCTGCTGACGACAGATCGACCTGCCACCCAC 436
 QY 859 CATCTCGTGTTCAGACGGGGAGGAGCCGCCACCTTACAGGGCCCTTGCACCTCCAGCT 918
 DB 437 CATCTCGTGTTCAGACGGGGAGGAGCCGCCACCTTACAGGGCCCTTGCACCTCCAGCT 496
 QY 919 TCGGACCCGAGCAGCAGCTGGAATCACTGAAACGGGAGTGTGTGGCGGACCCCAACAG 978
 DB 497 TCGGACCCGAGCAGCAGCTGGAATCACTGAAACGGGAGTGTGTGGCGGACCCCAACAG 556
 QY 979 AACCATCTTCGACAGTACCTGATGGATGTGCCAGGTGGGGCGGCCCTTCCGCCCCAG 1038
 DB 557 AACCATCTTCGACAGTACCTGATGGATGTGCCAGGTGGGGCGGCCCTTCCGCCCCAG 616
 QY 1039 CAGTAACTCGGGCATCAGGCCACGTGTATCGGAGCGCGGGCGGATGAGGGCGCGCC 1098
 DB 617 CAGTAACTCGGGCATCAGGCCACGTGTATCGGAGCGCGGGCGGATGAGGGCGCGCC 676
 QY 1099 GCGCACTTACAGCAGGTCATCGGCCACTTACCGGGGTCTCTTCAGGACACAGCAGAG 1158
 DB 677 GCGCACTTACAGCAGGTCATCGGCCACTTACCGGGGTCTCTTCAGGACACAGCAGAG 736
 QY 1159 CAGTGGGGCGCCCTCTTCTGTGGAGGAGCCCGGTCCACACACACATCGCGCCCT 1218

DB 737 CAGTGGGGCGCCCTCTTCTGTGGAGGAGCCCGGCTCCACACACATCGCGCCCT 796
 QY 1219 AGAGAGCGCAGCATCTCTGAGCAAGAGAGAGATAACAGAAAGCAGACACCTCTCTAGGG 1278
 DB 797 AGAGAGCGCAGCATCTCTGAGCAAGAGAGAGATAACAGAAAGCAGACACCTCTCTAGGG 856
 QY 1279 TCCCAGGGGGCGGGCTGGGGCTGCTAGGTGAAAAGGCGAG 1321
 DB 857 TCCCAGGGGGCGGGCTGGGGCTGCTAGGTGAAAAGGCGAG 899
 RESULT 6
 ADB75588
 ID ADB75588 standard; cDNA; 1141 BP.
 XX
 AC ADB75588;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Prostate cancer marker cDNA.
 XX
 KW Prostate; cancer; cytostatic; gene therapy; marker; ss.
 XX
 OS Homo sapiens.
 XX
 PN W02003009814-A2.
 XX
 PD 06-FEB-2003.
 XX
 PF 25-JUL-2002; 2002WO-US023913.
 XX
 PR 25-JUL-2001; 2001US-0307982P.
 PR 22-AUG-2001; 2001US-0314356P.
 PR 25-SEP-2001; 2001US-0325020P.
 PR 12-DEC-2001; 2001US-0341746P.
 PR 05-MAR-2002; 2002US-0362158P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbacheva B;
 PI Hoersh S, Kamatkar S, Wonsey AM, Glatt K, Zhao X, Anderson D;
 XX
 DR WPI; 2003-248033/24.
 XX
 PT New nucleic acid molecule, useful for diagnosing or treating prostate cancer.
 XX
 PS Disclosure; SEQ ID NO 412; 99pp; English.
 XX
 CC The invention relates to newly discovered cancer markers associated with the cancerous state of prostate cells. Also disclosed is a method of assessing whether a patient is afflicted with prostate cancer. The method of the invention involves assessing whether a patient is afflicted with prostate cancer by comparing the level of expression of a marker in a patient sample and the normal level of expression of the marker in a control non-prostate cancer sample, where a significant increase in the level of expression of the marker in the patient sample and the normal level indicates that the patient is afflicted with prostate cancer.
 CC Nucleic acids of the invention are useful for diagnosing or treating prostate cancer, and may be useful in gene therapy. Sequences given in ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1141 BP; 271 A; 350 C; 336 G; 184 T; 0 U; 0 Other;
 Query Match 60.7%; Score 802.2; DB 9; Length 1141;
 Best Local Similarity 98.4%; Pred. No. 3.9e-117;
 Matches 810; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
 QY 499 TTGTTCAGAGCTGGAGATCAGGAGCTGGAGTTTGTTCAGATCATCATCGTGGT 558

QY 1234 CTGGAGCAAGAGAGGATTAACAGAAAGGACACCTCTTAGGGTCCCAAGGGGGCGG 1293
 DB 722 CTGGAGCAAGAGAGGATTAACAGAAAGGACACCTCTTAGGGTCCCAAGGGGGCGG 781
 QY 1294 GCTGGGGCTGCTAGGTGAAAGGCAG 1321
 DB 782 GCTGGGGCTGCTAGGTGAAAGGCAG 809

RESULT 10

AAA75165
 ID AAA75165 standard, cDNA, 969 BP.
 AC AAA75165;
 XX
 DT 15-JAN-2001 (first entry)
 XX
 DE cDNA clone encoding a human TANGO 261 polypeptide.
 XX
 KW TANGO 266; TANGO 216; TANGO 261; TANGO 262; TANGO 267;
 KW cellular proliferation; cellular differentiation; cellular adhesion;
 KW von Willebrand factor-associated disorder; cell trafficking; cancer;
 KW hematopoietic associated disease; atelectasis; pulmonary congestion;
 KW edema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;
 KW intestinal disorder; spleen associated disease; renal disorder;
 KW cardiovascular disorder; ischemic heart disease; hydrocephalus;
 KW brain herniation; iatrogenic disease; inflammation; meningitis;
 KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;
 KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.
 XX Homo sapiens.

OS
 FH Key Location/Qualifiers
 FT CDS 6..764
 FT /*tag= a
 FT /product= "TANGO 261"

XX WO200052022-A1.
 PN 08-SEP-2000.
 XX 01-MAR-2000; 2000WO-US005226.
 XX 01-MAR-1999; 99US-0122458P.
 XX (MILL-) MILLENNIUM PHARM INC.
 PI Barnes TM, Holtzman DA, Sharp JD, Fraser CC;
 XX WPI: 2000-579269/54.
 DR P-PSDB; AAB18463.
 XX Novel human and murine secreted proteins designated TANGO 216, 261, 262,
 FT 266 and 267 useful as modulating agents of cellular processes, e.g. for
 PT treating cancer.
 PS Disclosure; Page; 175pp; English.
 XX
 CC AAA75163-65 encode human TANGO 261 proteins. The specification also
 CC describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO
 CC polypeptides can be used to modulate cellular proliferation, modulate
 CC cellular differentiation and/or modulate cellular adhesion. The proteins
 CC can be used to treat any von Willebrand factor-associated disorder,
 CC regulate extracellular matrix structuring, cellular adhesion, and cell
 CC trafficking and/or migration, modulate cellular interactions, modulate
 CC cell adhesion in proliferative disorders, such as cancer, modulate the
 CC proliferation, differentiation, and/or function of cells that appear in
 CC the bone marrow, and leukocytes, treat bone marrow, blood and
 CC hematopoietic associated diseases and disorders, atelectasis, pulmonary
 CC congestion or oedema, emphysema, chronic bronchitis, bronchial asthma and
 CC bronchiectasis, intestinal disorders, spleen associated diseases,
 CC modulate renal disorders, treat cardiovascular disorders such as ischemic

CC heart disease, modulate the proliferation, differentiation, and/or
 CC function of bone and cartilage cells and to treat bone and/or cartilage
 CC associated diseases or disorder. They may also be used to treat disorders
 CC associated with the ovaries, and cerebral oedema, hydrocephalus, brain
 CC herniations, iatrogenic disease, inflammations, bacterial and viral
 CC meningitis, Alzheimer's Disease, cerebral toxoplasmosis, Parkinson's
 CC disease, multiple sclerosis, brain cancers, hydrocephalus and
 CC encephalitis, and treat hepatic disorders. Note: the present sequence
 CC does not appear in the specification; it was created using information
 CC provided
 XX

SQ Sequence 969 BP; 210 A; 317 C; 294 G; 148 T; 0 U; 0 Other;
 Query Match 60.7%; Score 801.6; DB 3; Length 969;
 Best Local Similarity 99.5%; Pred. No. 4.9e-117;
 Matches 804; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 514 GGAGATCACGGAGCTGGAGTTTGTTCAGATCATCATCTCGTGGTGGTATGATGATGAT 573
 DB 2 GGAGATGGCGAGCTGGAGTTTGTTCAGATCATCATCTCGTGGTGGTATGATGATGAT 61
 QY 574 GGTGGTGGTATCACGTGCTCTGTAGCCACTACAGCTGTCTGCACGGTCTCTTCATCAG 633
 DB 62 GGTGGTGGTATCACGTGCTCTGTAGCCACTACAGCTGTCTGCACGGTCTCTTCATCAG 121
 QY 634 CCGGCACAGCCAGGGCGGAGAGAGAGATCCCTGTCTCAGAAAGGATGCTGTGGCC 693
 DB 122 CCGGCACAGCCAGGGCGGAGAGAGAGATCCCTGTCTCAGACGATGCTGTGGCC 181
 QY 694 CTCGGAGACACAGTGTCTAGGCAACGGAATCCAGAGCCGAGCTAGCCCGGCTCG 753
 DB 182 CTCGGAGACACAGTGTCTAGGCAACGGAATCCAGAGCCGAGCTAGCCCGGCTCG 241
 QY 754 GCCACACCGCTGGCCGTGCGCCCTTCGCCAGCGGAGCGCTTCCACCGCTTCCA 813
 DB 242 GCCACACCGCTGGCCGTGCGCCCTTCGCCAGCGGAGCGCTTCCACCGCTTCCA 301
 QY 814 GCCACCTATCGTACTCTGACGACAGATCGACCTGCCACCCACCATCTCGCTGTCAG 873
 DB 302 GCCACCTATCGTACTCTGACGACAGATCGACCTGCCACCCACCATCTCGCTGTCAG 361
 QY 874 CCGGAGAGGAGCCCGCCACCTACCGAGGGCCCTTGCACCTTCCAGCTTCCGAGACCC 933
 DB 362 CCGGAGAGGAGCCCGCCACCTACCGAGGGCCCTTGCACCTTCCAGCTTCCGAGACCC 421
 QY 934 GCAGCTGGAACTGAACCGGGAGTCCGTGCGCGGACCCCAACAGAGACCATCTTCGACAG 993
 DB 422 GCAGCTGGAACTGAACCGGGAGTCCGTGCGCGGACCCCAACAGAGACCATCTTCGACAG 481
 QY 994 TGACCTGATGGATAGTGCAGGCTGGGGCGGCCCTTGCACCTTCCAGCTTCCGAGCAT 1053
 DB 482 TGACCTGATGGATAGTGCAGGCTGGGGCGGCCCTTGCACCTTCCAGCTTCCGAGCAT 541
 QY 1054 CAGCGCACGTGTCTAGCGAGCGCGGGCGCATGAGGGGGCGCGCCCGCCACCTACAGCGA 1113
 DB 542 CAGCGCACGTGTCTAGCGAGCGCGGGCGCATGAGGGGGCGCGCCCGCCACCTACAGCGA 601
 QY 1114 GGTTCATCGGCCTACCCGGGGTCTCTTCAGAGACCCAGAGAGAGTGGCGGCCCTTC 1173
 DB 602 GGTTCATCGGCCTACCCGGGGTCTCTTCAGAGACCCAGAGAGAGTGGCGGCCCTTC 661
 QY 1174 CTTGCTGGAGGGACCCCGGCTCCACACACATATCGCGGCCCTTACAGAGCGAGCCAT 1233
 DB 662 CTTGCTGGAGGGACCCCGGCTCCACACACATATCGCGGCCCTTACAGAGCGAGCCAT 721
 QY 1234 CTTGAGCAAGAGAGAGATTAACAGAAAGGACACCTCTCTAGGGTCCCCAGGGGGCGG 1293
 DB 722 CTTGAGCAAGAGAGAGATTAACAGAAAGGACACCTCTCTAGGGTCCCCAGGGGGCGG 781
 QY 1294 GGTGGGGTCTAGGTGAAAGGCAG 1321
 DB 782 GGTGGGGTCTAGGTGAAAGGCAG 809

[illegible]

PN WO200034477-A2.
 XX 15-JUN-2000.
 PD 10-DEC-1999; 99WO-US030406.
 PF 11-DEC-1998; 98US-00210083.
 PR 09-FEB-1999; 99US-0119365P.
 PR 16-MAR-1999; 99US-0124687P.
 XX (INCY-) INCYTE PHARM INC.
 PA Tang YT, Yue H, Baughn MR, Hillman JL, Lal P, Au-Young J;
 PI Yang J, Lu DM, Azimzai Y;
 DR WPI; 2000-423423/36.
 DR P-PSDB; AAB01388.
 XX New human neuron-associated proteins and polynucleotides encoding them,
 PT useful for diagnosis, treatment and prevention of cell proliferative
 PT disorders including cancer, neuronal and neurological disorders.
 XX
 PS Claim 9; Page 136; 145pp; English.
 CC Human neuron-associated proteins (NEUAP) can be used for for treating or
 CC preventing a disorder associated with decreased expression or activity of
 CC NEUAP. Antagonists of NEUAP are useful for treating or preventing
 CC disorder associated with increased expression or activity of NEUAP. NEUAP
 CC or their fragments or derivatives are useful for treating neurological
 CC disorder such as epilepsy, ischemic cerebrovascular disease, stroke,
 CC cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's
 CC disease, dementia and Parkinson's disease. NEUAPs are also useful for
 CC treating other demyelinating diseases, bacterial and viral meningitis,
 CC prion diseases including kuru, Creutzfeldt-Jakob disease, nutritional and
 CC metabolic diseases of the nervous system, neurofibromatosis, other
 CC developmental disorders of the central nervous system, cerebral palsy,
 CC neuroskeletal disorders, autonomic nervous system disorders, cranial
 CC nerve disorders, spinal cord diseases, muscular dystrophy and other
 CC neuromuscular disorders, peripheral nervous system disorders, inherited,
 CC metabolic, endocrine, and toxic myopathies, mental disorders including
 CC mood, anxiety and schizophrenic disorders, a cell proliferative disorder
 CC such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis,
 CC cirrhosis, hepatitis, mixed connective tissue disease (MCTD),
 CC myelofibrosis, paroxysmal nocturnal haemoglobinuria, cancers of the
 CC adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, and an
 CC autoimmune/inflammatory disorder such as acquired immunodeficiency
 CC syndrome (AIDS), Addison's disease, adult respiratory distress syndrome,
 CC allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, Werner
 CC syndrome, complications of cancer, hemodialysis, and extracorporeal
 CC circulation, viral, bacterial, fungal parasitic, protozoal, and
 CC helminthic infections, and trauma. This sequence was given the incyte ID
 CC no. 1871288CB1
 XX
 SQ Sequence 1061 BP; 225 A; 342 C; 326 G; 168 T; 0 U; 0 Other;
 Query Match 60.6%; Score 800.6; DB 3; Length 1061;
 Best Local Similarity 98.3%; Pred. No 7e-117;
 Matches 809; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
 QY 499 TTTGTTCCAGAGATCGAGATCAGGAGTGGAGTTTGTTCAGATCATCATCGTGT 558
 DB 82 TCTCTCGGAAACAGGCAATCGGAGTGGAGTTTGTTCAGATCATCATCGTGT 141
 QY 559 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 618
 DB 142 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 201
 QY 619 ACSGTCTTCATCAGCGGCACAGCAGGCGGAGAGAGAGATGCGCTTCCTCAGA 678
 DB 202 ACGTCTTCATCAGCGGCACAGCAGGCGGAGAGAGAGATGCGCTTCCTCAGA 261
 QY 679 AGGATGCTGCTGCGCTCGAGAGCAGTGTTCAGGCAACGGAATCCAGAGCCCGAGT 738

DB 262 AGATGCTGTGCGCTCGAGAGACAGTGTTCAGGCAACGGAATCCAGAGCCCGAGT 321
 QY 739 CTACGCGCGCGCTCGCGCCACACCGCTGTGCGCTGTGCGCTGTGCGCTGTGCGCGAGCG 798
 DB 322 CTAGGCGCGCGCTCGCGCCACACCGCTGTGCGCTGTGCGCTGTGCGCGAGCG 381
 QY 799 CTTCCACCGCTTCAGCGCGCCACCTATCCGTACCTGACGACGACGAGATCGACCTGCCACCCAC 858
 DB 382 CTTCCACCGCTTCAGCGCGCCACCTATCCGTACCTGACGACGACGAGATCGACCTGCCCGCCAC 441
 QY 859 CATCTCGCTGTGAGCGGCGGAGCGCCACCTATCCAGGCGCGCTGTGACCTCCAGCT 918
 DB 442 CATCTCGCTGTGAGCGGCGGAGCGCCACCTATCCAGGCGCGCTGTGACCTCCAGCT 501
 QY 919 TCGGGACCGCGAGCAGCAGCTGGAACCTGAAACCGGAGTGGTGGCGCACCCCAACAG 978
 DB 502 TCGGGACCGCGAGCAGCAGCTGGAACCTGAAACCGGAGTGGTGGCGCACCCCAACAG 561
 QY 979 AACCATCTTCGACAGTGAACCTGATGATAGTCCAGGCTGGCGCGCTGTGCGCGCCAG 1038
 DB 562 AACCATCTTCGACAGTGAACCTGATGATAGTCCAGGCTGGCGCGCTGTGCGCGCCAG 621
 QY 1039 CAGTAACTCGGCTATCAGCGCGCACCTGTCTACGCGAGCGCGCGCGCTGTGCGCGCCAG 1098
 DB 622 CAGTAACTCGGCTATCAGCGCGCACCTGTCTACGCGAGCGCGCGCGCTGTGCGCGCCAG 681
 QY 1099 GCCCAGCTTACAGCGAGCTCATCGCGCACCTACCGCGGCTCTCTCTCCAGCAGCAGAG 1158
 DB 682 GCCCAGCTTACAGCGAGCTCATCGCGCACCTACCGCGGCTCTCTCTCCAGCAGCAGAG 741
 QY 1159 CAGTGGCGCGCTCTCTCTCTGAGGCGGAGCGCGCTCTCCAGCAGCAGCAGCAGCAGCAG 1218
 DB 742 CAGTGGCGCGCTCTCTCTCTGAGGCGGAGCGCGCTCTCCAGCAGCAGCAGCAGCAGCAG 801
 QY 1219 AGAGAGCGCAGCATCTGAGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1278
 DB 802 AGAGAGCGCAGCATCTGAGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 861
 QY 1279 TCCCGCAGGCGCGCGCTGTGCGCTGTGCGCTGTGCGCTGTGCGCTGTGCGCTGTGCG 1321
 DB 862 TCCCGCAGGCGCGCGCTGTGCGCTGTGCGCTGTGCGCTGTGCGCTGTGCGCTGTGCG 904
 RESULT 13
 ADC37324
 ID ADC37324 standard; DNA; 1085 BP.
 AC ADC37324;
 XX 18-DEC-2003 (first entry)
 DT Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 157.
 DE
 XX Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
 KW cancer; infectious disease; bone disease; AIDS;
 KW neurodegenerative disease; ischaemic disorder; Antiinflammatory;
 KW Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-Hiv;
 KW Neuroprotective; Nootropic; Cardiant; Gene therapy; human; gene; ds.
 XX
 OS Homo sapiens.
 XX
 XX WO2003048202-A2.
 XX
 XX 12-JUN-2003.
 XX
 XX 03-DEC-2002; 2002WO-JP012644.
 XX
 XX 03-DEC-2001; 2001JP-00368692.
 PR 05-DEC-2001; 2001US-0335829P.
 PR 03-OCT-2002; 2002JP-00291302.
 PR 04-OCT-2002; 2002US-0415769P.
 XX
 XX (ASAH) ASAH KASEI KK.
 PA

XX Matsuda A, Muramatsu S;
 XX WPI: 2003-505282/47.
 DR P-PSDB; ADC37325.
 XX
 PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
 PT useful for treating inflammation, autoimmune diseases, cancers,
 PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
 PT ischemic disorders.
 XX
 XX Claim 4; SEQ ID NO 157; 938pp; English.
 PS
 CC The present invention relates to novel proteins and their coding
 CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
 CC kappaB). The proteins and their coding sequences are useful for treating
 CC a disease associated with NF-kappaB activation, such as inflammation,
 CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
 CC neurodegenerative diseases, or ischaemic disorders.
 XX
 SQ Sequence 1085 BP; 223 A; 352 C; 334 G; 176 T; 0 U; 0 Other;
 Query Match 60.6%; Score 800.6; DB 9; Length 1085;
 Best Local Similarity 98.3%; Pred. No. 7e-117;
 Matches 809; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
 QY 499 TTGTTCCAGCATGAGATCAGATCAGGAGCTGGAGTTTGTTCAGATCATCATCTGCTG 558
 DB 109 TCTCTCGAAACACGCAATGGCGGAGCTGGAGTTTGTTCAGATCATCATCTGCTG 168
 QY 559 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 618
 DB 169 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 228
 QY 619 ACGTCTCTTATCAGCGGCACACCGCAGGCGGAGAGAGAGATGCCCTGTCTCAGA 678
 DB 229 ACGTCTCTTATCAGCGGCACACCGCAGGCGGAGAGAGAGATGCCCTGTCTCAGA 288
 QY 679 AGGATGCTGTGGCCCTCGAGAGCAGATGTCTAGGCAACGGAATCCAGACCGCAGGT 738
 DB 289 AGGATGCTGTGGCCCTCGAGAGCAGATGTCTAGGCAACGGAATCCAGACCGCAGGT 348
 QY 739 CTACGCGCGGCTCGGCGCCACCGACCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCG 798
 DB 349 CTACGCGCGGCTCGGCGCCACCGACCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCG 408
 QY 799 CTTCACCGCTTCAGCGCCACTATCGTACCTGACGACGATGACCTGCCACCCAC 858
 DB 409 CTTCACCGCTTCAGCGCCACTATCGTACCTGACGACGATGACCTGCCACCCAC 468
 QY 859 CATCTCGTGTTCAGACGGGAGGAGCCGCCACCTACAGGCGCCCTGTCCACCTCCAGCT 918
 DB 469 CATCTCGTGTTCAGACGGGAGGAGCCGCCACCTACAGGCGCCCTGTCCACCTCCAGCT 528
 QY 919 TCGGACCCCGCAGCAGCAGCTGGAACCTGAACCGGAGTGGTGGCGGCGACCCCAACAG 978
 DB 529 TCGGACCCCGCAGCAGCAGCTGGAACCTGAACCGGAGTGGTGGCGGCGACCCCAACAG 588
 QY 979 AACCATCTTCAGACAGTACCTGATGATGATGATGATGATGATGATGATGATGATGAT 1038
 DB 589 AACCATCTTCAGACAGTACCTGATGATGATGATGATGATGATGATGATGATGATGAT 648
 QY 1039 CAGTAACCTCGGCGCATCAGCGCCAGTGTGTACGGCAGCGCGGCGGCGATGAGGGCGGCC 1098
 DB 649 CAGTAACCTCGGCGCATCAGCGCCAGTGTGTACGGCAGCGCGGCGGCGATGAGGGCGGCC 708
 QY 1099 GCCCACTTCAGACAGGAGTTCATCGCCACTACCGCGGGTCTCTCTTCAGACACAGCAGAG 1158
 DB 709 GCCCACTTCAGACAGGAGTTCATCGCCACTACCGCGGGTCTCTCTTCAGACACAGCAGAG 768
 QY 1159 CAGTGGCGCGCT 1218
 DB 769 CAGTGGCGCGCT 828

QY 1219 AGAGAGCGCAGCATCTGGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1278
 DB 829 AGAGAGCGCAGCATCTGGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 888
 QY 1279 TCCCGAGCGGCGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 1321
 DB 889 TCCCGAGCGGCGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 931
 RESULT 14
 ABZ36103
 ID ABZ36103 standard; cDNA; 1334 BP.
 XX
 AC ABZ36103;
 DT 10-FEB-2003 (first entry)
 DE Human secretory polynucleotide SPTM SEQ ID NO 267.
 XX
 KW Human; SPTM; autoimmune disorder; inflammatory disorder; AIDS; anaemia;
 KW asthma; Crohn's disease; neurological disorder; epilepsy; cancer;
 KW Huntington's disease; Alzheimer's disease; Creutzfeldt-Jakob disease;
 KW multiple sclerosis; Parkinson's disease; cell proliferative disorder;
 KW anti-inflammatory; immunosuppressive; neuroprotective; nontropic;
 KW neuroleptic; anticonvulsant; cytostatic; antiparkinsonian; anxiolytic;
 KW antipsoriatic; antianaemic; anti-HIV; human immunodeficiency virus;
 KW secretory polynucleotide; secretory protein; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200283876-A2.
 XX
 PD 24-OCT-2002.
 XX
 PF 27-MAR-2002; 2002WO-US009921.
 XX
 PR 29-MAR-2001; 2001US-0280067P.
 PR 29-MAR-2001; 2001US-0280068P.
 PR 16-MAY-2001; 2001US-0291280P.
 PR 17-MAY-2001; 2001US-0291829P.
 PR 17-MAY-2001; 2001US-0291849P.
 PR 19-JUN-2001; 2001US-0299428P.
 PR 20-JUN-2001; 2001US-029976P.
 PR 20-JUN-2001; 2001US-030001P.
 XX
 (INCY-) INCYTE GENOMICS INC.
 XX
 PA Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
 PI Dufford GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshey SR;
 PI Daughtery SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;
 PI Peralta CH, David KH, Lewis SA, Chen AJ, Panzer SR, Harris B;
 PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
 XX
 DR WPI: 2003-075543/07.
 DR P-PSDB; ABP75860.
 XX
 PT New human secretory proteins and polynucleotides, useful for diagnosing,
 PT treating or preventing autoimmune/inflammatory disorders (e.g. AIDS),
 PT neurological disorders (e.g. Alzheimer's), or cell proliferations or
 PT cancers.
 XX
 PS Claim 1; SEQ ID NO 267; 459pp + Sequence Listing; English.
 XX
 CC The invention relates to a secretory polynucleotide (designated sptm)
 CC comprising any of 567 polynucleotide sequences (ABZ35837-ABZ36403), a
 CC naturally occurring polynucleotide sequence at least 90 % identical to
 CC the polynucleotide sequence, a polynucleotide complementary to them or an
 CC RNA equivalent of them. The polypeptide or polynucleotide are useful for
 CC treating, preventing or diagnosing a disease or condition associated with
 CC the expression of functional SPTM. These are particularly useful for
 CC diagnosing, treating or preventing autoimmune/inflammatory disorders
 CC (e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's

CC disease), neurological disorders (e.g. epilepsy, Huntington's disease,
CC dementia, stroke, Alzheimer's disease, Creutzfeldt-Jakob disease,
CC multiple sclerosis, cerebral palsy, Parkinson's disease, anxiety,
CC schizophrenia or amnesia), or cell proliferative disorders (e.g.
CC psoriasis, polycythemia vera, or cancers including adenocarcinoma,
CC leukaemia, lymphoma, melanoma, myeloma, sarcoma or cancers of the brain,
CC breast, cervix or prostate). Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1334 BP; 299 A; 381 C; 394 G; 260 T; 0 U; 0 Other;
Query Match 60.6%; Score 800.6; DB 7; Length 1334;
Best Local Similarity 98.3%; Pred. No. 6.9e-117;
Matches 809; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 499 TTGTTCCAGACATGAGATCAGCGAGCTGGAGTTTGTTCAGATCATCATCGTGGT 558
DB 82 TCTCTCGGAAACCAAGGCAATGGCGAGCTGGAGTTTGTTCAGATCATCATCGTGGT 141
QY 559 GGTGATGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 618
DB 142 GGTGATGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 201
QY 619 ACGTTCCTTCATAGCGGCAAGCCAGCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 678
DB 202 ACGTTCCTTCATAGCGGCAAGCCAGCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 261
QY 679 AGGATGCTGTGCGCTCGAGAGACAGATGTCAGGCAACGGAATCCAGAGCGCGAGGT 738
DB 262 AGGATGCTGTGCGCTCGAGAGACAGATGTCAGGCAACGGAATCCAGAGCGCGAGGT 321
QY 739 CTACGCGCGCTTCGCGCCACCGAGCGCGCTGGCGCGCTGGCGCGCTGGCGCGCTGGCG 798
DB 322 CTACGCGCGCTTCGCGCCACCGAGCGCGCTGGCGCGCTGGCGCGCTGGCGCGCTGGCG 381
QY 799 CTTCCACCGCTTCCAGCGCCACCTATCCGTACCTGAGCAGCAGATCGACCTGCCACCCAC 858
DB 382 CTTCCACCGCTTCCAGCGCCACCTATCCGTACCTGAGCAGCAGATCGACCTGCCACCCAC 441
QY 859 CATCTCGCTGTGAGAGCGGAGAGCGCCACCTTACAGGCGCGCTGACCGCTCCAGCT 918
DB 442 CATCTCGCTGTGAGAGCGGAGAGCGCCACCTTACAGGCGCGCTGACCGCTCCAGCT 501
QY 919 TCGGAGCCCGAGCAGCAGCTGGAATGAACCGGAGTCTGGTGGCGGACCCCAAGACAG 978
DB 502 TCGGAGCCCGAGCAGCAGCTGGAATGAACCGGAGTCTGGTGGCGGACCCCAAGACAG 561
QY 979 AACCATCTTCGACAGTGACCTGATGATAGTGCAGGCTGGCGCGCTGGCGCGCTGGCG 1038
DB 562 AACCATCTTCGACAGTGACCTGATGATAGTGCAGGCTGGCGCGCTGGCGCGCTGGCG 621
QY 1039 CAGTAATCGGGATCAGCGCCACGTGCTACCGCAGCGCGCGCGCGCGCGCGCGCGCGCG 1098
DB 622 CAGTAATCGGGATCAGCGCCACGTGCTACCGCAGCGCGCGCGCGCGCGCGCGCGCGCG 681
QY 1099 GCCCACTACAGGAGGTGATCGGCCACTACCGGGGTCTCTCTCCAGCAGCAGCAGAG 1158
DB 682 GCCCACTACAGGAGGTGATCGGCCACTACCGGGGTCTCTCTCCAGCAGCAGCAGAG 741
QY 1159 CAGTGGCGCGCTCTCTCTGAGGAGGACCGCGTCCACACACACATCGCGCGCT 1218
DB 742 CAGTGGCGCGCTCTCTCTGAGGAGGACCGCGTCCACACACACATCGCGCGCT 801
QY 1219 AGAGAGCGCGATCTGGAGCAAGAGAGAGATTAACAGAGGACCCCTCTCTAGGG 1278
DB 802 AGAGAGCGCGATCTGGAGCAAGAGAGAGATTAACAGAGGACCCCTCTCTAGGG 861
QY 1279 TCCCCAGGGGCGCGCGCTGGCGCTCGTAGTGTAAGGACAG 1321
DB 862 TCCCCAGGGGCGCGCGCTGGCGCTCGTAGTGTAAGGACAG 904

RESULT 15
AA157868
ID AA157868 standard; cDNA; 1066 BP.
XX
AC AA157868;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 71.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00489725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
PA (HYSE-) HYSEQ INC.
XX
Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX
WPI; 2001-442253/47.
DR P-PSDB; AAM38712.
XX
Novel nucleic acids and polypeptides, useful for treating disorders such
as central nervous system injuries.
XX
Claim 1; SEQ ID NO 71; 10078pp; English.
XX
The invention relates to human nucleic acids (AA157798-AA161369) and the
encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
XX part of the printed specification
SQ Sequence 1066 BP; 225 A; 348 C; 325 G; 168 T; 0 U; 0 Other;
Query Match 59.8%; Score 790.4; DB 4; Length 1066;
Best Local Similarity 99.2%; Pred. No. 2.8e-115;
Matches 794; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 522 CGGAGCTGGAGTTTGTTCAGATCATCATCGTGGTGGTGGTGGTGGTGGTGGTGGTGG 581

Db	113	CGGACGACGATCTTGTTAGATCATCATCGTGGTGATGATGGTGGTGG	172
Qy	582	TGATCAGCTGCTGTGAGCCACTACAGCTGTGTGCACGGTCTTTCATCAGCCGGCACA	641
Db	173	TGATCAGCTGCTGTGAGCCACTACAGCTGTGTGCACGGTCTTTCATCAGCCGGCACA	232
Qy	642	GCCAGGGCGGAGGAGAGAGATGCCCTGTCTCTCAGAAAGATGCTGTGGCCCTCGGAGA	701
Db	233	GCCAGGGCGGAGGAGAGAGATGCCCTGTCTCTCAGAAAGATGCTGTGGCCCTCGGAGA	292
Qy	702	GCACAGTGTACGACAAAGAAATCCAGAGCGGAGGTCTACGCCCGGCTCGGCCACCG	761
Db	293	GCACAGTGTACGACAAAGAAATCCAGAGCGGAGGTCTACGCCCGGCTCGGCCACCG	352
Qy	762	ACCGCTGCGCGCTGCGCCCTTCGCCAGCGGAGGCTTCACACGCTTCCAGGCCACCT	821
Db	353	ACCGCTGCGCGCTGCGCCCTTCGCCAGCGGAGGCTTCACACGCTTCCAGGCCACCT	412
Qy	822	ATCCGTACCTGACGACGAGATCGACTGCTCCACCCACCATCTCGTGTCTCAGACGGGAGG	881
Db	413	ATCCGTACCTGACGACGAGATCGACTGCTCCACCCACCATCTCGTGTCTCAGACGGGAGG	472
Qy	882	AGCCCCACCTACCGAGGCGCTTCGACCTTCAGCTTCGGACCCCGAGCAGCAGCTGG	941
Db	473	AGCCCCACCTACCGAGGCGCTTCGACCTTCAGCTTCGGACCCCGAGCAGCAGCTGG	532
Qy	942	AACTGAACCGGAGTGGTGGCGCACCCGCCAAACAGAACCATCTTCGACAGTGACCTGA	1001
Db	533	AACTGAACCGGAGTGGTGGCGCACCCGCCAAACAGAACCATCTTCGACAGTGACCTGA	592
Qy	1002	TGGATAGTGCAGGCTGGGGCGCCCTGCCCCGAGCAGTAACTCGGGCATCAGGCCA	1061
Db	593	TGGATAGTGCAGGCTGGGGCGCCCTGCCCCGAGCAGTAACTCGGGCATCAGGCCA	652
Qy	1062	CGTGCTACGGCAGCGGGCGCATGGAGGGGCGCGCCACCTACAGCGAGGTCTATCG	1121
Db	653	CGTGCTACGGCAGCGGGCGCATGGAGGGGCGCGCCACCTACAGCGAGGTCTATCG	712
Qy	1122	GCCACTACCGGGTCTCTTCAGACACAGCAGAGTGGCGCCCTCTCTGCTGG	1181
Db	713	GCCACTACCGGGTCTCTTCAGACACAGCAGAGTGGCGCCCTCTCTGCTGG	772
Qy	1182	AGGGGACCCGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCATCTGGAGCA	1241
Db	773	AGGGGACCCGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCATCTGGAGCA	832
Qy	1242	AAGAGAGGATAAAGAGAGACACCTCTCTAGGGTCCCTAGGGGGCGGGCTGGGG	1301
Db	833	AAGAGAGGATAAAGAGAGACACCTCTCTAGGGTCCCTAGGGGGCGGGCTGGGG	892
Qy	1302	CTGGTAGGTGAAAGGCAG	1321
Db	893	CTGGTAGGTGAAAGGCAG	912

Search completed: May 12, 2004, 00:56:42
Job time : 606.354 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 11, 2004, 21:10:37 ; Search time 5450.49 Seconds
(without alignments)
10504.776 Million cell updates/sec

Title: US-09-934-249-1

Perfect score: 1321
Sequence: 1 cgaccgcggtctcggagcga.....ctgcgtagtgaaaggcag 1321

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:

- 1: gb_ba:*
- 2: gb_btg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: gb_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rod:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1321	100.0	1321	6	AX392417 Sequence
2	1229	93.0	4839	9	AF305616 Homo sapi
3	1227.4	92.9	1383	6	AX775889 Sequence
4	887.4	67.2	1061	9	BC015918 Homo sapi
5	861	65.2	861	6	AX392419 Sequence
6	803.2	60.8	969	6	BD272494 Secreted
7	802.2	60.7	1141	9	AF224278 Homo sapi
8	801.8	60.7	1913	6	BD272544 Secreted
9	801.6	60.7	969	6	BD272514 Secreted
10	801.6	60.7	969	6	BD272515 Secreted
11	801.6	60.7	969	6	BD272516 Secreted
12	801.6	60.7	1140	6	AX336830 Sequence
13	800.6	60.6	1085	6	AX775887 Sequence
14	800	60.6	1818	9	AX128643 Homo sapi
15	755.8	57.2	759	6	AX336831 Sequence
16	755.4	57.2	759	6	BD272545 Secreted
17	751.2	56.9	756	6	BD272495 Secreted
18	750.4	56.8	753	6	BD272534 Secreted
19	630	47.7	1379	10	BC036995 Mus muscu
20	616.4	46.7	878	6	AX392428 Sequence
21	594.4	45.0	61505	9	AF305426 Homo sapi
22	594.4	45.0	130435	9	HS71807 Human DNA
23	589.2	44.6	1583	6	AX593655 Sequence
24	522.8	39.6	150224	9	HSJ1059L7 Human DNA
25	493.6	37.4	693	6	AX392430 Sequence
26	466.2	35.3	1713	6	BD272504 Secreted
27	464.6	35.2	1713	6	BD272517 Secreted
28	464.6	35.2	1713	6	BD272518 Secreted
29	464.6	35.2	1713	6	BD272519 Secreted
30	461.2	34.9	651	6	BD272505 Secreted
31	458.2	34.7	648	6	BD272535 Secreted
32	425	32.2	651	10	AF220208 Mus muscu
33	417.8	31.6	812	6	AX011709 Sequence
34	417.8	31.6	812	6	BD226320 Pancreat
35	401.4	30.4	408	6	AX071267 Sequence
36	400.6	30.3	673	6	AX525744 Sequence
37	377.2	28.6	156698	10	AL837509 Mouse DNA
38	377.2	28.6	175754	2	AC110189 Mus muscu
39	377.2	28.6	231930	2	AC134911 Mus muscu
40	357.6	27.1	249554	2	AC139417 Rattus no
41	357.6	27.1	258632	2	AC111878 Rattus no
42	352.2	26.7	921	6	AR233384 Sequence
43	352.2	26.7	8065	6	AR233383 Sequence
44	352.2	26.7	8494	9	AF009424 Homo sapi
45	350	26.5	2170	6	AX713513 Sequence

ALIGNMENTS

RESULT 1

AX392417

LOCUS

DERIVATION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

AX392417 Sequence 1 from Patent WO0216416.
AX392417 Sequence 1 from Patent WO0216416.
AX392417.1 GI:19700732
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Lee,R.T., Landschulz,K.T., Kennedy,S.P., Thompson,J.F. and
Turi,T.O.
Diagnosis and treatment of cardiovascular conditions

Pred. No. is the number of results predicted by chance to have a

JOURNAL Patent: WO 0216416-A 1 28-FEB-2002;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US) ; PFIZER INC. (US)
FEATURES Location/Qualifiers
source 1..1321
CDS /organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
413..1276
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD29005.1"
/db_xref="GI:19700733"
/translation="MRLMGVNSTAAAGAPNVSCNCKRSIFQSWETEFVQI
IIIVVMVWVITLLSHVKLSARFISRSOGRREDALSEGCLWPSEBTVSGN
GTPBQVAVPREPTDLAVPPFAOPRPVYLOHEIDLPPTISLDGEPPEP
YQGPCTLOLRPEOLENRESVRAPNRTIEDSLMDSARLGPCPPSSNSGISATC
YSGGRMGPPPTSEVIGHYPGSGFQHQSSGPPSLLEGRLRHHHTIAPLESAAIWS
KENDKQKQGHPL"

ORIGIN
Query Match 100.0%; Score 1321; DB 6; Length 1321;
Best Local Similarity 100.0%; Pred. No. 1.6e-176;
Matches 1321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACCGCGTCTCGAGAGAAACCGATCTCTTGGACTTGAATGAGAGGAGGCGG 60
DB 1 CGACCGCGTCTCGAGAGAAACCGATCTCTTGGACTTGAATGAGAGGAGGCGG 60
QY 61 CGGCGCGCGCGCGGAGGCGCTCGGCTGGGGAAGCTAGCGGACAGGCTCAGCCCC 120
DB 61 CGGCGCGCGCGCGGAGGCGCTCGGCTGGGGAAGCTAGCGGACAGGCTCAGCCCC 120
QY 121 GCGCGAGCGCGCGCGCGCTCGGAGCGCAATTTTCGAGCGCCACCCCGGCGCATCGCG 180
DB 121 GCGCGAGCGCGCGCGCGCTCGGAGCGCAATTTTCGAGCGCCACCCCGGCGCATCGCG 180
QY 181 AGCGCGCGCGCGCTCGGAGGAGCGCGGCGGCGCGGAGCGGCGGCGCGGCGGCG 240
DB 181 AGCGCGCGCGCGCTCGGAGGAGCGCGGCGGCGCGGAGCGGCGGCGGCGGCGGCG 240
QY 241 TGAGCGCGCGCGCGCGCGGAACTTGGCGGCGACCCGAGCGCGCGCGCGCGCGCGCG 300
DB 241 TGAGCGCGCGCGCGCGCGGAACTTGGCGGCGACCCGAGCGCGCGCGCGCGCGCGCG 300
QY 301 CTCCCGCGCGCGCGCTCTCGATCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
DB 301 CTCCCGCGCGCGCGCTCTCGATCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
QY 361 CCCCG 420
DB 361 CCCCG 420
QY 421 CTGTGAGGGGTCAACAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
DB 421 CTGTGAGGGGTCAACAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
QY 481 GTGCAACTCAAAACGCTCTTTCTTCAGAGCATGAGATCACGAGCTGGAGTTGTTC 540
DB 481 GTGCAACTCAAAACGCTCTTTCTTCAGAGCATGAGATCACGAGCTGGAGTTGTTC 540
QY 541 GATCATCATCTGCTGGTGGTGAATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 541 GATCATCATCTGCTGGTGGTGAATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 601 CCACTACAAGCTGTCTGCAAGCTCTTCTATACGCGCGCACAGCGAGGCGGAGAGAGA 660
DB 601 CCACTACAAGCTGTCTGCAAGCTCTTCTATACGCGCGCACAGCGAGGCGGAGAGAGA 660
QY 661 AGATGCCCTGTCTCAGAAAGATGCTGTGCGCCCTCGGAGAGCACAGTGTCAAGCAAG 720
DB 661 AGATGCCCTGTCTCAGAAAGATGCTGTGCGCCCTCGGAGAGCACAGTGTCAAGCAAG 720

QY 721 AATCCAGAGCGCGAGGTCTAGCCCGCGCTCGSCCCACCGACCGCTGGCGCTGCCGCC 780
DB 721 AATCCAGAGCGCGAGGTCTAGCCCGCGCTCGSCCCACCGACCGCTGGCGCTGCCGCC 780
QY 781 CTTGCGCCAGCGGAGCGCTTCCACCGCTTCCAGCGCCACCTATCGTACCTGACAGCA 840
DB 781 CTTGCGCCAGCGGAGCGCTTCCACCGCTTCCAGCGCCACCTATCGTACCTGACAGCA 840
QY 841 GATCGACCTGCGCACACCATCTCGCTGTACAGCGGGAGGAGCCCCACCTACACAGG 900
DB 841 GATCGACCTGCGCACACCATCTCGCTGTACAGCGGGAGGAGCCCCACCTACACAGG 900
QY 901 CCCCTGCACCTTCCAGCTTTCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGAGT 960
DB 901 CCCCTGCACCTTCCAGCTTTCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGAGT 960
QY 961 GCGCGACCCCCAAAACAGAACCATCTTTCACAGTACCTGATGGATAGTGCAGGCTGG 1020
DB 961 GCGCGACCCCCAAAACAGAACCATCTTTCACAGTACCTGATGGATAGTGCAGGCTGG 1020
QY 1021 CGGCGCTGCCCCCGCCAGCAGTAACCTCGGGCATCAGCGCCACCTGCTACGCGAGCGCG 1080
DB 1021 CGGCGCTGCCCCCGCCAGCAGTAACCTCGGGCATCAGCGCCACCTGCTACGCGAGCGCG 1080
QY 1081 GCGCATGAGAGGCGCGCGCGCCACCTACAGCAGGTCATCGGCGCTACCCGGGTCCTC 1140
DB 1081 GCGCATGAGAGGCGCGCGCGCCACCTACAGCAGGTCATCGGCGCTACCCGGGTCCTC 1140
QY 1141 CTTCCAGCACCAGCAGCAGTGGGCGCGCTTCTTGTGGAGGGAGCCCGCTCCACCA 1200
DB 1141 CTTCCAGCACCAGCAGCAGTGGGCGCGCTTCTTGTGGAGGGAGCCCGCTCCACCA 1200
QY 1201 CACACATCGCGCGCTTAGAGAGCGCAGCCATCTTGGAGCAAGAGAGATTAACAGAA 1260
DB 1201 CACACATCGCGCGCTTAGAGAGCGCAGCCATCTTGGAGCAAGAGAGATTAACAGAA 1260
QY 1261 AGGACACCTCTCTAGGCTCCCGAGGGGCGCGGCTGGGCTGCTAGTGAAGAGCA 1320
DB 1261 AGGACACCTCTCTAGGCTCCCGAGGGGCGCGGCTGGGCTGCTAGTGAAGAGCA 1320
QY 1321 G 1321
DB 1321 G 1321

RESULT 2
AF305616
LOCUS Homo sapiens STAG1/PMEP1 mRNA, complete cds.
DEFINITION AF305616
ACCESSION AF305616
VERSION AF305616.1 GI:16303741
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4839)
AUTHORS Rae, F.K., Hooper, J.D., Nicol, D.L. and Clements, J.A.
TITLE Characterization of a novel gene, STAG1/PMEP1, up-regulated in
renal cell carcinoma and other solid tumors
JOURNAL Mol. Carcinog. 32 (1), 44-55 (2001)
MEDLINE 21453682
PUBMED 11568975
REFERENCE 2 (bases 1 to 4839)
AUTHORS Rae, F.K., Hooper, J.D., Nicol, D.L. and Clements, J.A.
TITLE Direct Submission
JOURNAL Submitted (14-SEP-2000) Centre for Molecular Biotechnology,
Queensland University of Technology, 2 George St, Brisbane, QLD
4001, Australia
FEATURES Location/Qualifiers
source 1..4839
/organism="Homo sapiens"
/mol_type="mRNA"

QY	93	GGAAAGCTAGCGCAGAGGCTCAGCCCCCGCGCGCAGCGCGCCCGCTGCCAGGCCATT	152
DB	1	GGAAAGCTAGCGCAGAGGCTCAGCCCCCGCGCGCAGCGCGCCCGCTGCCAGGCCATT	60
QY	153	TTCCGAGACCAACCCCGGGCACTGCGAGCGCCCCCGGGGCTGCGCAGGGGAGCGCGGG	212
DB	61	TTCCGAGACCAACCCCGGGCACTGCGAGCGCCCCCGGGGCTGCGCAGGGGAGCGCGGG	120
QY	213	GGCGCAGCGGAGCGGGTCCC CGCAGCTAGAGCCCCCGCGCGCCCCGGGAATTTGCGCG	272
DB	121	GGCGCAGCGGAGCGGGTCCC CGCAGCTAGAGCCCCCGCGCGCCCCGGGAATTTGCGCG	180
QY	273	GACCGAGCCGCGGAGCGGGCGCGCTCCCGCGCGCGCGCTCCGCTGCAATGCGGGG	332
DB	181	GACCGAGCCGCGGAGCGGGCGCGCTCCCGCGCGCGCGCTCCGCTGCAATGCGGGG	240
QY	333	CCAGCTCCGGGCGCGGCGCGGCGCGCCCCCGCGCGCGCCCCCGAGCCCCCGCGCCCCG	392
DB	241	CCAGCTCCGGGCGCGGCGCGGCGCGCCCCCGCGCGCGCCCCCGAGCCCCCGCGCCCCG	300
QY	393	GCGCGCGCGCGGCGCGTCCATGCACCGCTTGATGGGGTCAACAGCAGCGCGCGCGCG	452
DB	301	GCGCGCGCGCGGCGCGTCCATGCACCGCTTGATGGGGTCAACAGCAGCGCGCGCGCG	360
QY	453	CCGCGGGCAGCCCAATGTCTCTGCACTGCAACTGCAAAAGCTCTTTGTTCCAGAGCA	512
DB	361	CCGCGGGCAGCCCAATGTCTCTGCACTGCAACTGCAAAAGCTCTTTGTTCCAGAGCA	420
QY	513	TGAGATCAAGGAGCTGGAGTTGTCAGATCATCATCTGCTGGTGGTGGTGGTGGTGA	572
DB	421	TGAGATCAAGGAGCTGGAGTTGTCAGATCATCATCTGCTGGTGGTGGTGGTGGTGA	480
QY	573	TGTTGGTGGTGATCACCTGCTGCTGAGCCACTACAAGCTGTGTGACGCTCTTCATCA	632
DB	481	TGTTGGTGGTGATCACCTGCTGCTGAGCCACTACAAGCTGTGTGACGCTCTTCATCA	540
QY	633	GCGGCAAGCAGCGGCGGAGGAGAGAAATGCCCTGTCTCTCAGAGAGTGCCTGGC	692
DB	541	GCGGCAAGCAGCGGCGGAGGAGAGAAATGCCCTGTCTCTCAGAGAGTGCCTGGC	600
QY	693	CCTCGGAGACAGATGTCAAGCAACGGAAATCCAGAGCGCGAGTCTACGCCCGCCCTC	752
DB	601	CCTCGGAGACAGATGTCAAGCAACGGAAATCCAGAGCGCGAGTCTACGCCCGCCCTC	660
QY	753	GGCCACCGACCGCTGGCGGTGGCGCTTTGGCCAGCGGAGCGCTTCCACCGCTTC	812
DB	661	GGCCACCGACCGCTGGCGGTGGCGCTTTGGCCAGCGGAGCGCTTCCACCGCTTC	720
QY	813	AGCCACCTATCCGTACTCTGAGCAGAGATCGACCTGCCACCCACCATCTCGCTGTCA	872
DB	721	AGCCACCTATCCGTACTCTGAGCAGAGATCGACCTGCCACCCACCATCTCGCTGTCA	780
QY	873	ACGGGAGGAGCCCCCAACCTTACAGGGCCCCCTGCAACCTTCAGCTTCGGGACCCGAGC	932
DB	781	ACGGGAGGAGCCCCCAACCTTACAGGGCCCCCTGCAACCTTCAGCTTCGGGACCCGAGC	840
QY	933	AGCAGCTGGAACTGAAACGGGAGTCGGTGGCGCACCCCAACAGAAACCATCTTCACA	992
DB	841	AGCAGCTGGAACTGAAACGGGAGTCGGTGGCGCACCCCAACAGAAACCATCTTCACA	900
QY	993	GTGACCTGATGGATAGTGCCAGGTGGCGGCCCCCTGCCCCCCCAGCAGTAATTCGGGCA	1052
DB	901	GTGACCTGATGGATAGTGCCAGGTGGCGGCCCCCTGCCCCCCCAGCAGTAATTCGGGCA	960
QY	1053	TCAGCGCCACCTGTACGGCAGCGGCGGCGCATGGAGGGCGCGCCCCCACCCTACAGCG	1112
DB	961	TCAGCGCCACCTGTACGGCAGCGGCGGCGCATGGAGGGCGCGCCCCCACCCTACAGCG	1020
QY	1113	AGGTCAATCGGCACTAACCCGGGTCTCTCTTCCAGCAACAGCAGAGCAGTGGGCGCCCT	1172
DB	1021	AGGTCAATCGGCACTAACCCGGGTCTCTCTTCCAGCAACAGCAGAGCAGTGGGCGCCCT	1080

Qy	1173	CTTGTGTGAGGGGACCCGGCTCCACACACACATCGGCCCTCTAGAGAGCGCAGCCA	1232
Db	1081	CTTGTGTGAGGGGACCCGGCTCCACACACACATCGGCCCTCTAGAGAGCGCAGCCA	1140
Qy	1233	TCTGGAGCAAGAAGATGATTAACAGAAAGGACACCTCTCTAGGCTCCCCAGGGGGGCC	1292
Db	1141	TCTGGAGCAAGAAGATGATTAACAGAAAGGACACCTCTCTAGGCTCCCCAGGGGGGCC	1200
Qy	1293	GGGCTGGGGCTGCTAGGTGAAAAGGCAG	1321
Db	1201	GGGCTGGGGCTGCTAGGTGAAAAGGCAG	1229

RESULT 4	BC015918	1061 bp	mRNA	linear	PRI 07-OCT-2003
LOCUS	BC015918				
DEFINITION	Homo sapiens transmembrane, prostate androgen induced RNA, mRNA (CDNA clone MGC:20374 IMAGE:4559576), complete cds.				
ACCESSION	BC015918				
VERSION	BC015918.1	GI:16198474			
KEYWORDS	MGC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 1061)				
AUTHORS	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Dege,J.G., Klausner,R.D., Collins,F.S., Wagner,K.H., Shermen,C.M., Schuler,G.D., Altshul,S.F., Zeeberg,B., Euetow,K., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,D., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,I., Stapietom,M., Soares,M.B., Donald,M.F., Usdin,T.B., Toshiyuki,S., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Loquellano,N.A., Peters,G.J., Carninci,P., Prange,C., Raha,S.S., Logue,W.N., McEwan,P.J., Abramson,R.D., Mullah,S.J., Bosak,S.A., Richards,S., McKernan,K.J., Malek,J.A., Gunatratne,P.H., Richardson,S., Worley,K.C., Hale,S.G., Garcia,A.M., Gay,L.J., Hulyk,S.W., Wallaon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahy,J., Heiton,E., Kettelman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalls,D.E., Schnarch,A., Schein,J.E., Jones,S.J. and Marra,M.A.				
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
MEDLINE	22388257				
PUBMED	12477932				
REFERENCE	2 (bases 1 to 1061)				
AUTHORS	Strausberg,R.				
TITLE	Direct Submission				
JOURNAL	Submitted (15-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
REMARK	NTH-MGC Project URL: http://mgc.nci.nih.gov				
COMMENT	Contact: MGC help desk Email: cgabs-rc@mail.nih.gov Tissue Procurement: DCTD/DTT cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@bcgsc.bc.ca Steven Jones, Jennifer Amano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Lilja Prabhur, Parvaneh Saedi, Jacqueline Schnell, Duane Smalls, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasha van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.				

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 29 Row: e Column: 5
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

FEATURES
SOURCE

```

location:Qualifiers
1..1061
/organism="Homo sapiens"
/mol_type="mRNA"
/db xref="taxon:9606"
/cisone="MGC:20374 IMAGE:4559576"
/cisone_type="Kidney, renal cell adenocarcinoma"
/cisone_lib="NH MGC,14"
/lab host="DH10B-R"
/note="Vector: pOTB7"
1..1061
/gene="TMEPAI"
/note="synonym: PMEPAI"
/db xref="LOCUS:56937"
/db xref="MIM:606564"
131..844
/codon_start=1
/product="TMEPAI protein"
/protein_id="AAH15918.1"
/db xref="GI:16198475"
/db xref="Locustid:56937"
/translations="MMVMVVITCLSHYKLSARSFISRHSQGRREDALSSSGCLMP
SESTVGGNGPEQVYAPRRDRLAIVPPFAQRHFRHFQTYTYLQGHIDLPTTSL
SDGEVPPGPGTQLGRDEQQLNRESVRAPRNTIFQSLMDLSARLQGCPCPSS
NSGISTATSCSGMGMEGPPPTTYVEIGHYGPSSFOHQQSGPPSLLELRLHHTHTAP
LEGAATWSKEDKQKGKGP"

```

ORIGIN

Query Match	67.23;	Score 887.4;	DB 9;	Length 1061;
Best Local Similarity	99.94;	Pred. No. 1.6e-115;		
Matches 888;	Conservative 0;	Mismatches 11;	Indels 0;	Gaps 0;
QY	433	CAACAGCACCGCCGCGCCGCGCGGGCAGCCCAATGTCTCTCTGCACGTGCACACTGC	CAA	492
Db	1	CAACAGCACCGCCGCGCCGCGCGGGCAGCCCAATGTCTCTCTGCACGTGCACACTGC	CAA	60
QY	493	ACGCTCTTTGTTCCAGAGCATGGAGATCACGGAGCTGGAGTTTGTTCCAGATCATCAT	TCAT	552
Db	61	ACGCTCTTTGTTCCAGAGCATGGAGATCACGGAGCTGGAGTTTGTTCCAGATCATCAT	TCAT	120
QY	553	CGTGGTGGTGATGATGGTGATGGTGGTGATCACGTGCGTCTGAGGCCACTACAA	AGCT	612
Db	121	CGTGGTGGTGATGATGGTGATGGTGGTGATCACGTGCGTCTGAGGCCACTACAA	AGCT	180
QY	613	GTCTGCACGGTCTTCATCAGCCGGCACAGCCAGGGCGGAGGAGAGAAATGCCCT	GTCTGC	672
Db	181	GTCTGCACGGTCTTCATCAGCCGGCACAGCCAGGGCGGAGGAGAGAAATGCCCT	GTCTGC	240
QY	673	CTCAGAGGATGCTGTGGCCCTCGGAGAGCACAGTGTCAAGCAACGGAATCCCA	GAGCC	732
Db	241	CTCAGAGGATGCTGTGGCCCTCGGAGAGCACAGTGTCAAGCAACGGAATCCCA	GAGCC	300
QY	733	GCAGGTCTAGCCCCGCTCGGCCACCGACCGGCTGGCGTGGCGCCCTTCGCC	ACG	792
Db	301	GCAGGTCTAGCCCCGCTCGGCCACCGACCGGCTGGCGTGGCGCCCTTCGCC	ACG	360
QY	793	GGAGCGTCTCCACCGCTTCCAGGCCACCTATCCGTAGCTGCAGCACGAGATCGA	CTGCC	852
Db	361	GGAGCGTCTCCACCGCTTCCAGGCCACCTATCCGTAGCTGCAGCACGAGATCGA	CTGCC	420
QY	853	ACCCACCATCTCGTGTTCAGACGGGAGGAGCCCCACCTTACACAGGGCCCTTC	GCACCT	912
Db	421	GCCACCATCTCGTGTTCAGACGGGAGGAGCCCCACCTTACACAGGGCCCTTC	GCACCT	480
QY	913	CCAGCTTCGGGAACCCGAGCAGCAGCTGGAACTGAACCGGAGTTCGGTGGCGCAC	CCCC	972

Db	481	CGAGCTTCGGGACCCCGAGCAGCAGCTGGNACTGAACCCGGAGTGCCTGGTGCGCGCACCCCC	540	
Qy	973	AAACAGAACCATCTTCGACAGTAGTCACTGATGTAGATAGTGCACGAGCTGGGCGGCCCTTCGCC	1032	
Db	541	AAAACAGAACCATCTTCGACAGTAGTCACTGATGTAGTGCACGAGCTGGGCGGCCCTTCGCC	600	
Qy	1033	CCCAGCAGTAATCTCGGGCATCAGCGCCACGCTCTACGGCAGCGCGGGCGCATGGAGGG	1092	
Db	501	CCCAGCAGTAATCTCGGGCATCAGCGCCACGCTCTACGGCAGCGCGGGCGCATGGAGGG	660	
Qy	1093	GCGCGCGGCCCATCTCAGCAGGAGTCACTCGGCCACTACCOCGGGTCTCTCTTCAGCACCA	1152	
Db	661	GCGCGCGGCCCATCTCAGCAGGAGTCACTCGGCCACTACCOCGGGTCTCTCTTCAGCACCA	720	
Qy	1153	GCAGAGCAGTGGGCGGCCCTCTTTGCTGGAGGGGACCCGGCTCCACACACACACATCGC	1212	
Db	721	GCAGAGCAGTGGGCGGCCCTCTTTGCTGGAGGGGACCCGGCTCCACACACACATCGC	780	
Qy	1213	GCCCTTAGAGAGCGCAGCCACTCTGGAGCAAGAGAGGATAAACAAGAGGACACCTCT	1272	
Db	781	GCCCTTAGAGAGCGCAGCCACTCTGGAGCAAGAGAGGATAAACAAGAGGACACCTCT	840	
Qy	1273	CTAGGGTCCCCACGGGGCGCGGCTCGGGCTCGTAGGTGAAAAGCGAG	1321	
Db	841	CTAGGGTCCCCACGGGGCGCGGCTCGGGCTCGTAGGTGAAAAGCGAG	889	
RESULT	5			
AX392419				
LOCUS	AX392419	861 bp DNA linear	PAT 23-MAR-2002	
DEFINITION	Sequence 3 from Patent WO0216416.			
ACCESSION	AX392419			
VERSION	AX392419.1	GI:19700734		
KEYWORDS	Homo sapiens (human)			
SOURCE	Homo sapiens			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1	Lee, R.T., Landschulz, K.T., Kennedy, S.P., Thompson, J.F. and Turi, T.G.		
AUTHORS				
TITLE	Diagnosis and treatment of cardiovascular conditions			
JOURNAL	Patent: WO 0216416-A 3 28-FEB-2002;			
	THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US); PFIZER INC. (US)			
FEATURES	Location/Qualifiers			
source	1..861	/notes="unnamed protein product"		
	/organism="Homo sapiens"	/codon_start=1		
CDS	/mol_type="unassigned DNA"	/protein_id="CAD29006.1"		
	/db_xref="taxon:9606"	/db_xref="GI:19700735"		
	1..>861	/translation="MERLGVNSTAAAAAGQPNVSTCNKRSLPFOSMEITELFEVQI IIIVVMVMVVITCLLSHYKLARSFISRHSGRRREDALSSGCLMPSSESTVGSN GIPEOVPAWPTDLAVHPAQRRFRFPQTYPLQHEIDLPPTISLSDGEZPPP YQPCFTQLRDPDQOLAKNESVRAPNPRTIFDSLMDARLGGPPCPSPNSIGSATC VGSNGRMGPMPPTYSEVIGHYPGSSSFHQHQQSSGPPSLLEGTRLHHHTIAPLESAAIYS KEKDQKQHPL"		
ORIGIN				
Query Match	65.2%	Score 861; DB 6; Length 861;		
Best Local Similarity	100.0%;	Pred. No. 8.4e-112;		
Matches	861; Conservative	0; Mismatches	0; Indels	
		0; Gaps	0;	
Qy	413	ATGCACGCTTGATGGGGTCAACACACACCGCGCGCGCGCGCGGAGCCCAATGTC	472	
Db	1	ATGCACGCTTGATGGGGTCAACACACACCGCGCGCGCGCGCGGAGCCCAATGTC	60	
Qy	473	TCTGCACGTGCAACTGCAAAACGCTCTTTTCCAGAGCATGGAGATCACGAGCTGGAG	532	

QY 1294 GGCTGGGGCTGGCTAGTGAAGAGGCGAG 1321
 DB 782 GGCTGGGGCTGGCTAGTGAAGAGGCGAG 809

RESULT 7
 AF224278 1141 bp mRNA linear PRI 18-JUL-2000
 LOCUS Homo sapiens PMEPA1 protein (PMEPA1) mRNA, complete cds.
 DEFINITION AF224278
 ACCESSION AF224278.1 GI:9255808
 VERSION
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1141)
 Xu, L.L., Shanmugam, N., Segawa, T., Sesterhenn, I.A., McLeod, D.G.,
 Moul, J.W., and Srivastava, S.
 A novel androgen-regulated gene, PMEPA1, located on chromosome
 20q13 exhibits high level expression in prostate
 Genomics 66 (3), 257-263 (2000)

TITLE
 20q13 exhibits high level expression in prostate

JOURNAL
 MEDLINE
 PUBMED
 10873380

REFERENCE
 2 (bases 1 to 1141)
 Xu, L.L., Shanmugam, N., Segawa, T., Sesterhenn, I.A., McLeod, D.G.,
 Moul, J.W., and Srivastava, S.
 Direct Submission
 Submitted (12-JAN-2000) CPDR, USUHS, 1530 East Jefferson Street,
 Rockville, MD 20852, USA

FEATURES
 source
 1. .1141
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="20"
 /map="20q13.31-13.33"
 /cell_line="LNCaP"
 1. .1141
 /gene="PMEPA1"
 96. .854
 /gene="PMEPA1"
 /note="type 1b transmembrane protein; expression is
 induced in response to the synthetic androgen, R1881;
 expression is abundant in, and restricted to, prostate
 glandular epithelial cells; similar to the predicted
 protein encoded by sequence deposited at GenBank Accession
 Number NP_004329"
 /codon_start=1
 /product="PMEPA1 protein"
 /protein_id="AAF86322.1"
 /db_xref="GI:9255809"
 /translations="MALEFVQIIIVVMVMMVVITLGHVLSARSFISRHSGQ
 RRREDALSSGCLWPSESTVSGNGIPEPOVYAPRPTDLAVPPFAORFRFRFOPT
 PYLQHEIDLPTTISLSDGERPPYQGPCTQLDPEQQLERLRESVRAPNRTIFDS
 LMSRLRGCPSSNSGISATCYGSGRMGPPYISVIGHYPOSSFHQSSGPP
 SLLEGTRLRHTHIAPEUSAIAWSKDKQKHPL"

ORIGIN
 Query Match 60.7%; Score 802.2; DB 9; Length 1141;
 Best Local Similarity 98.4%; Pred. No. 1.5e-103;
 Matches 810; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 499 TTCTGTCAGAGCATGAGATCAGCGAGCTGGAGTTTGTTCAGATCATCATCTCGTGGT 558
 DB 77 TCTCCTCGAAACACAGCAATGGCGGAGCTGGAGTTTGTTCAGATCATCATCTCGTGGT 136

QY 559 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 618
 DB 137 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 196

QY 519 ACGGTCTCTTATCATGCCCGGCACACCCAGGCGGAGAGAGAGATGCCCTGTCTCAGAG 678

DB 197 ACGGTCTCTTATCATGCCCGGCACAGCCAGGGCGGAGAGAGAGATGCCCTGTCTCAGAG 256
 QY 679 AGGATGCTCTGTGGCCCTCGAGAGACACAGTGTCTAGGCAACCGCAATCCAGAGCGCGCAGGT 738
 DB 257 AGGATGCTCTGTGGCCCTCGAGAGACACAGTGTCTAGGCAACCGCAATCCAGAGCGCGCAGGT 316
 QY 739 CTACGCCCGGCTCGGCCCAACGACCGGCTGGCGCTGGCGCCCTTCGCCCAAGCGGAGCG 798
 DB 317 CTACGCCCGGCTCGGCCCAACGACCGGCTGGCGCTGGCGCCCTTCGCCCAAGCGGAGCG 376
 QY 799 CTTCCACCGCTTCAGAGCCACCTATTCGTACTCTCAGCAGCAGAGATCGACCTGCCACCCAC 858
 DB 377 CTTCCACCGCTTCAGAGCCACCTATTCGTACTCTCAGCAGCAGAGATCGACCTGCCACCCAC 436
 QY 859 CATCTCGTGTCTCAGAGCGGAGAGAGCCACCTACCTACAGGGCCCTGTGCACCTCTCCAGCT 918
 DB 437 CATCTCGTGTCTCAGAGCGGAGAGAGCCACCTACCTACAGGGCCCTGTGCACCTCTCCAGCT 496
 QY 919 TCGGAGCCCGCAGCAGCAGCTGGAACCTGAACCGGGAGTTCGTGCGCGCACCCCAACAG 978
 DB 497 TCGGAGCCCGCAGCAGCAGCTGGAACCTGAACCGGGAGTTCGTGCGCGCACCCCAACAG 556
 QY 979 AACCATCTTCACAGTGTACCTGTAGTGTGAGTGTGCGGCGGCGCCCTCTGCCCCCGCAG 1038
 DB 557 AACCATCTTCACAGTGTACCTGTAGTGTGAGTGTGCGGCGGCGCCCTCTGCCCCCGCAG 616
 QY 1039 CAGTAACTTCGGGCTACAGCGGAGAGTGTCTAGCGGAGCGGCGGCGGCGCATGAGGGGCGCC 1098
 DB 617 CAGTAACTTCGGGCTACAGCGGAGAGTGTCTAGCGGAGCGGCGGCGGCGCATGAGGGGCGCC 676
 QY 1099 GCGCACCTTACAGCGAGGTTCATCGGCGCATACCGCGGGTTCCTCTTCAGCAGCACCAGCAG 1158
 DB 677 GCGCACCTTACAGCGAGGTTCATCGGCGCATACCGCGGGTTCCTCTTCAGCAGCACCAGCAG 736
 QY 1159 CAGTGGGCGGCGCTCTCTGTCTGAGGGAGACCGCGCTCCACACACACATCGGCGCGCT 1218
 DB 737 CAGTGGGCGGCGCTCTCTGTCTGAGGGAGACCGCGCTCCACACACACATCGGCGCGCT 796
 QY 1219 AGAGAGCGGCGCATCTCGAGCAAGAGAGATAAACAGAAAGACACCTCTCTTAGGG 1278
 DB 797 AGAGAGCGGCGCATCTCGAGCAAGAGAGATAAACAGAAAGACACCTCTCTTAGGG 856
 QY 1279 TCCCGAGGGGCGGCGGCTGGGCTGGGTGTAGGTGAAGAGCGAG 1321
 DB 857 TCCCGAGGGGCGGCGGCTGGGCTGGGTGTAGGTGAAGAGCGAG 899

RESULT 8
 BD272544 1913 bp DNA linear PAT 17-JUL-2003
 LOCUS Secreted proteins and nucleic acids encoding them.
 DEFINITION BD272544
 ACCESSION BD272544
 VERSION BD272544.1 GI:33082312
 KEYWORDS JP 2002539773-A/53.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1913)
 Barnes, T.M., Holtzman, D.A., Sharp, J.D. and Fraser, C.C.
 Secreted proteins and nucleic acids encoding them
 Patent: JP 2002539773-A 53 26-NOV-2002;
 JOURNAL MILLENNIUM PHARMACEUTICALS INC
 COMMENT OS Homo sapiens (human)
 FN JP 2002539773-A/53
 PD 26-NOV-2002
 PF 01-MAR-2000 JP 2000602247
 PR 01-MAR-1999 US 60/122458
 PI THOMAS M BARNES, DOUGLAS A HOLTZMAN, JOHN D SHARP, CHRISTOPHER C
 FR FRASER
 PC C12N15/09, C07K14/47, C07K16/18, C12N5/10, C12P21/02, C12Q1/68, PC
 G01N33/15,


```

Db      482  TGACCTGATGGATAGTGCAGGCTGGCGGCGCCCTGCCCCCAGCAGTAACCTCGGGCAT 541
Qy      1054  CAGGGCCACCTGCTACCGGACGGCGCGGCATGAGGGGCCCGCCACCTACAGCGA 1113
Db      542  CAGCGCCACCTGCTACCGGACGGCGCGGCATGAGGGGCCCGCCACCTACAGCGA 601
Qy      1114  GGTCAATCGGCACTACCGGCGGCTCTCTTCCAGCACACAGAGAGTGGGCGGCCCTC 1173
Db      602  GGTCAATCGGCACTACCGGCGGCTCTCTTCCAGCACACAGAGAGTGGGCGGCCCTC 661
Qy      1174  CTTGCTGGAGGGACCGGCTCTACACACACACATCGCCGCCCTGAGAGCGACCCAT 1233
Db      662  CTTGCTGGAGGGACCGGCTCTACACACACACATCGCCGCCCTGAGAGCGACCCAT 721
Qy      1234  CTGAGCAAGAGAGAGAGATAAACAGAAAGACACCCCTCTAGGGTCCCCAGGGGGCGG 1293
Db      722  CTGAGCAAGAGAGAGATAAACAGAAAGAGACACCCCTCTAGGGTCCCCAGGGGGCGG 781
Qy      1294  GGCTGGGGCTCGTAGGTGAAAAGGCAG 1321
Db      782  GGCTGGGGCTCGTAGGTGAAAAGGCAG 809

RESULT 10
BD272515 LOCUS 969 bp DNA linear PAT 17-JUL-2003
DEFINITION Secreted proteins and nucleic acids encoding them.
ACCESSION BD272515
VERSION BD272515.1 GI:33082283
KEYWORDS JP 2002539773-A/24.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Barnes,T.M., Holtzman,D.A., Sharp,J.D. and Fraser,C.C.
TITLE Secreted proteins and nucleic acids encoding them
JOURNAL Patent: JP 2002539773-A 24 26-NOV-2002;
MILLENNIUM PHARMACEUTICALS INC
COMMENT OS Homo sapiens (human)
PN JP 2002539773-A/24
PC 26-NOV-2002
PF 01-MAR-2000 JP 200602247
PR 01-MAR-1999 US 60/122458
PI THOMAS M BARNES, DOUGLAS A HOLTZMAN, JOHN D SHARP, CHRISTOPHER C
PI FRASER
PC C12N15/09,C07K14/47,C07K16/18,C12N5/10,C12P21/02,C12Q1/68, PC
G01N33/15,
G01N33/50,G01N33/53,G01N33/53,G01N33/566,C12N15/00,C12N5/00 CC
Secreted proteins and nucleic acids encoding them FH Key
Location/Qualifiers
FT CDS Location/Qualifiers
source 1..969
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 60.7%; Score 801.6; DB 6; Length 969;
Best Local Similarity 99.5%; Pred. No. 1.9e-103;
Matches 804; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      514  GGAGATCAGGAGCTGGAGTTGTTTCAGATCATCATCATCGTGGTGGTGGTGGTGGT 573
Db      2  GGAGATGGCGAGCTGGAGTTGTTTCAGATCATCATCATCGTGGTGGTGGTGGTGGT 61
Qy      574  GGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 633
Db      62  GGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 121
Qy      634  CCGGCACACAGCCAGGGGGGAGAGAGAGATGCCCTGTCTCAGAGAGATGCCCTGTG 693

```

```

Db      122  CCGGCACAGCCAGGGGGGAGAGAGAGATGCCCTGTCTCAGAGAGATGCCCTGTGCC 181
Qy      694  CTCGAGAGAGACAGTGTTCAGCAACGGAATCCAGAGCGCGAGGTCTACGGCGCGCTCG 753
Db      182  CTCGAGAGAGACAGTGTTCAGCAACGGAATCCAGAGCGCGAGGTCTACGGCGCGCTCG 241
Qy      754  GCCACACCGCCCTGTGGCGGCTCTTCCAGCACAGAGAGAGGTCTTCCACCGCTTCCA 813
Db      242  GCCACACCGCCCTGTGGCGGCTCTTCCAGCACAGAGAGAGGTCTTCCACCGCTTCCA 301
Qy      814  GCCACCTATCTGATGATCTGAGCAGAGAGATGCACTGCCACCCACCTCTGCTGTGAGA 873
Db      302  GCCACCTATCTGATGATCTGAGCAGAGAGATGCACTGCCACCCACCTCTGCTGTGAGA 361
Qy      874  CCGGAGAGAGCGCCCGCCCTTACAGGGCCCTGCAACCTCCAGGTTCGGGACCCCGAGCA 933
Db      362  CCGGAGAGAGCGCCCGCCCTTACAGGGCCCTGCAACCTCCAGGTTCGGGACCCCGAGCA 421
Qy      934  GCAGCTGGAACTGAAACCGGGAGTGGTGGCGGCAACCCCAACAGAACCATCTTTCGACAG 993
Db      422  GCAGCTGGAACTGAAACCGGGAGTGGTGGCGGCAACCCCAACAGAACCATCTTTCGACAG 481
Qy      994  TGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1053
Db      482  TGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 541
Qy      1054  CAGCGCCACGCTGCTACGCGAGCGGCGGCGCATGAGAGGGCGCGGCCACCTACAGCGA 1113
Db      542  CAGCGCCACGCTGCTACGCGAGCGGCGGCGCATGAGAGGGCGCGGCCACCTACAGCGA 601
Qy      1114  GGTCAATCGGCACTACCGGCGGCTCTCTTCCAGCACACAGAGAGTGGGCGGCCCTC 1173
Db      602  GGTCAATCGGCACTACCGGCGGCTCTCTTCCAGCACACAGAGAGTGGGCGGCCCTC 661
Qy      1174  CTTGCTGGAGGGACCGCGGCTCCACACACACATCGCGGCCCTTAGAGAGCGACCCAT 1233
Db      662  CTTGCTGGAGGGACCGCGGCTCCACACACACATCGCGGCCCTTAGAGAGCGACCCAT 721
Qy      1234  CTTGAGCAAGAGAGAGATAAACAGAAAGAGACCCCTCTAGGGTCCCCAGGGGGCGG 1293
Db      722  CTTGAGCAAGAGAGAGATAAACAGAAAGAGACCCCTCTAGGGTCCCCAGGGGGCGG 781
Qy      1294  GGCTGGGGCTCGTAGGTGAAAAGGCAG 1321
Db      782  GGCTGGGGCTCGTAGGTGAAAAGGCAG 809

RESULT 11
BD272516 LOCUS 969 bp DNA linear PAT 17-JUL-2003
DEFINITION Secreted proteins and nucleic acids encoding them.
ACCESSION BD272516
VERSION BD272516.1 GI:33082284
KEYWORDS JP 2002539773-A/25.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Barnes,T.M., Holtzman,D.A., Sharp,J.D. and Fraser,C.C.
TITLE Secreted proteins and nucleic acids encoding them
JOURNAL Patent: JP 2002539773-A 25 26-NOV-2002;
MILLENNIUM PHARMACEUTICALS INC
COMMENT OS Homo sapiens (human)
PN JP 2002539773-A/25
PD 26-NOV-2002
PF 01-MAR-2000 JP 200602247
PR 01-MAR-1999 US 60/122458
PI THOMAS M BARNES, DOUGLAS A HOLTZMAN, JOHN D SHARP, CHRISTOPHER C
PI FRASER
PC C12N15/09,C07K14/47,C07K16/18,C12N5/10,C12P21/02,C12Q1/68, PC
G01N33/15,

```

```

PC   G01N33/50,G01N33/53,G01N33/53,G01N33/566,C12N15/00,C12N5/00 CC
Secreted proteins and nucleic acids encoding them FH   Key
Location/Qualifiers
FT   CDS           (5)..(761)..
      Location/Qualifiers
      1..969
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
FEATURES
      source

```

ORIGIN

[illegible]

RESULT 12

```

Db      811  CTGAGCAAGAAAGGATTAACAGAAAGACACCTCTCTAGGTGCCACAGGGGGCG 870
QY      1294  GGCTGGGGCTGCGTAGGTGAAAAGGCAG 1321
Db      871  GGCTGGGGCTGCGTAGGTGAAAAGGCAG 898

RESULT 13
AX775887
LOCUS      AX775887      1085 bp      mRNA      linear      PAT 14-JUL-2003
DEFINITION Sequence 157 from Patent WO03048202.
ACCESSION AX775887
VERSION   AX775887.1 GI:32693605
KEYWORDS  Homo sapiens (human)
SOURCE    Homo sapiens
ORGANISM  Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Matsuda, A. and Muramatsu, S.
NF-kB activating gene
Patent: WO 03048202-A 157 12-JUN-2003;
Asahi Kasei Kabushiki Kaisha (JP)
Location/Qualifiers
FEATURES
source     1..1085
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            128..886
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="CA511643.1"
            /db_xref="GI:32693606"
            /translation="MAELEFVQIIIVMMVAVITCLSHYKLSABSF1SRHSQ
            RREDAISSGCLWPSESTVSGNGIFBPQVAPRETDLAVPPFAQRPRFPQPT
            FYLOHEIDLPTISLSGERPPYQGPCTQLQRPDPEQLNRESVAPRNPRTIFDSD
            LMSARLGFCPPSSNSGIGATCYGSGRMGPPYFSEVIGHYPOSSFOHQSSGPP
            SLLEGTRLHHTHIAPLESAAIWSKDKQKHPL"

ORIGIN
Query Match      60.6%; Score 800.5; DB 6; Length 1085;
Best Local Similarity 98.3%; Pred No. 2.5e-103;
Matches 809; Conservative 0; Mismatches 14; Indels 0; Caps 0;

QY      499  TTTGTTCCAGAGCAGGATCAGCGAGCTGGAGTTGTTTCAGATCATCATCGTGGT 558
Db      109  TCTCTGCGAAACACGAGCAATGCGGAGCTGGAGTTGTTTCAGATCATCATCGTGGT 168
QY      559  GGTGATGATGTGATGTGTGTGATCAGTGTGCTGCTGAGCCACTACAGCTGTGC 618
Db      169  GGTGATGATGTGATGTGTGTGATCAGTGTGCTGCTGAGCCACTACAGCTGTGC 228
QY      619  ACGGTCTTTCATCAGCGGCAACGAGCGGCGGAGGAGAGATGCCCTGTCTCAGA 678
Db      229  ACGGTCTTTCATCAGCGGCAACGAGCGGCGGAGGAGAGATGCCCTGTCTCAGA 288
QY      679  AGGATGCTGTGGCTCGGAGAGACAGTGTGAGGCAACGGAATCCAGACGGCAGGT 738
Db      289  AGGATGCTGTGGCTCGGAGAGACAGTGTGAGGCAACGGAATCCAGACGGCAGGT 348
QY      739  CTACGCCCCGCTCGGCGCAACGAGCGCTGCGCGCTGCGCGCTGCGCGCGGAGCG 798
Db      349  CTACGCCCCGCTCGGCGCAACGAGCGCTGCGCGCTGCGCGCTGCGCGCGGAGCG 408
QY      799  TTTCACCGCTTCAGCGGAGGAGCCCTATCCGTAACGAGACAGATGACCTGCCAGCT 858
Db      409  TTTCACCGCTTCAGCGGAGGAGCCCTATCCGTAACGAGACAGATGACCTGCCAGCT 468
QY      859  CATCTCGTGTGAGCGGAGGAGGAGCCCTATCCGTAACGAGGAGGAGGAGGAGGAG 918
Db      459  CATCTCGTGTGAGCGGAGGAGGAGGAGCCCTATCCGTAACGAGGAGGAGGAGGAG 528
QY      919  TCGGAGCCCGAGCAGCAGCTGGAACTGAACCGGAGTCGCTGCGCGACCCCAACAG 978

```

```

Db      529  TCAGGACCCCGAGCAGAGCTGGAACCGCGAGTGGGTGCGCGACCCCAACAG 588
QY      979  AACCATCTTCGACAGTGCCTGATGATGATGATGATGATGATGATGATGATGATGAT 1038
Db      589  AACCATCTTCGACAGTGCCTGATGATGATGATGATGATGATGATGATGATGATGAT 648
QY      1039  CAGTAATCTCGGCGATCAGCGCCACGTCGTCAGGAGCGGGCGGCGGCGGCGGCGGCC 1098
Db      649  CAGTAATCTCGGCGATCAGCGCCACGTCGTCAGGAGCGGGCGGCGGCGGCGGCGGCC 708
QY      1099  GCCCACCTACAGCAGGAGTGCATCGGCCACTACCCCGGGTCTCTCTCCAGACACACAG 1158
Db      709  GCCCACCTACAGCAGGAGTGCATCGGCCACTACCCCGGGTCTCTCTCCAGACACACAG 768
QY      1159  CAGTGGGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1218
Db      769  CAGTGGGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 828
QY      1219  AGAGAGCGCAGCATCTCTGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1278
Db      829  AGAGAGCGCAGCATCTCTGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 888
QY      1279  TCCCGAGCGGCGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 1321
Db      889  TCCCGAGCGGCGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 931

RESULT 14
AX128643
LOCUS      AX128643      1818 bp      mRNA      linear      PRI 10-APR-2003
DEFINITION Homo sapiens PMEPAL variant A protein mRNA, complete cds.
ACCESSION AX128643
VERSION   AX128643.1 GI:22121998
KEYWORDS  Homo sapiens (human)
SOURCE    Homo sapiens
ORGANISM  Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1818)
  Brunschwig, E.B., Wilson, K., Mack, D., Dawson, D., Lawrence, E.,
  Willson, J.K.V., Lu, S., Nosrati, A., Rerko, R.M., Swinler, S., Beard, L.,
  Lutterbaugh, J.D., Willis, J., Platzner, P. and Markowitz, S.
  PMEPAL, a transforming growth factor-beta-induced marker of
  terminal colonocyte differentiation whose expression is maintained
  in primary and metastatic colon cancer
  Cancer Res. 63 (7), 1568-1575 (2003)
  22557253
  MEDLINE
  PUBMED
REFERENCE
1 (bases 1 to 1818)
  Brunschwig, E.B., Wilson, K., Mack, D., Dawson, D., Lawrence, E.,
  Willson, J.K.V., Lu, S., Nosrati, A., Swinler, S., Beard, L.,
  Lutterbaugh, J.D., Willis, J., Platzner, P. and Markowitz, S.
  Direct Submission
  Submitted (03-JUL-2002) Department of Medicine, Case Western
  Reserve University/Howard Hughes Medical Institute, 11001 Cedar
  Ave., Cleveland, OH 44106, USA
  Location/Qualifiers
  1..1818
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /chromosome="20"
    /map="20q13.31-13.33"
    145..858
    /codon_start=1
    /product="PMEPAL variant A protein"
    /protein_id="AA089277.1"
    /db_xref="GI:22121998"
    /translation="MMVAVITCLSHYKLSARFISRHSGRRRDLSSGCLMP
    SSVSTVGIGIPBPQVAPRETDLAVPPFAQRERFHPQPTFYLOHEIDLPTISL
    SGEPPYQGPCTQLQRPDPEQLNRESVAPRNPRTIFDSDMSARLGFCPPSS
    NSGIGATCYGSGRMGEGPPYFSEVIGHYFGSSFOHQSSGPPSLLEGTRLHHTIAP

CDS

```


Db 331 GCGCCACGACCGCTGGCGCTGCGCCCTTGGCCAGCGGAGCGCTTCCACCGCTTCCA 390
Qy 814 GCGCCACCTATCGTACCTGACACAGATCGACCTGCGACCGCAGCATCTCGTGTGAGA 873
Db 391 GCGCCACCTATCGTACCTGACACAGATCGACCTGCGACCGCAGCATCTCGTGTGAGA 450
Qy 874 CGGGAGAGGAGCCCAACCTTACAGGGCCCTGACCTTCAAGCTTGGGAGCCCGAGCA 933
Db 451 CGGGAGAGGAGCCCAACCTTACAGGGCCCTGACCTTCAAGCTTGGGAGCCCGAGCA 510
Qy 934 GCAGCTGGAACTGAACCGGAGTGGTGGCGCACCCCAACAGAACCATCTTCGACAG 993
Db 511 GCAGCTGGAACTGAACCGGAGTGGTGGCGCACCCCAACAGAACCATCTTCGACAG 570
Qy 994 TGACCTGATGATAGTGCAGCTGGCGGCGCCCTTGGCCCGCCCGCAGCATCTGGGAT 1053
Db 571 TGACCTGATGATAGTGCAGCTGGCGGCGCCCTTGGCCCGCCCGCAGCATCTGGGAT 630
Qy 1054 CAGCGCCAGTCTACGCGCAGCGCGGCGCATGAGGGCGCGCCCGCCAGCATCTGAGCGA 1113
Db 631 CAGCGCCAGTCTACGCGCAGCGCGGCGCATGAGGGCGCGCCCGCCAGCATCTGAGCGA 690
Qy 1114 GGTATCGGCCACTACCGGGTCTCTCTTCCAGCACCGACAGCAGTGGGCGCCCTC 1173
Db 691 GGTATCGGCCACTACCGGGTCTCTCTTCCAGCACCGACAGCAGTGGGCGCCCTC 750
Qy 1174 CTGCTGGAGGGAGCCCGCTCCACACACACATCGCCCTTAGAGGGCGAGCCAT 1233
Db 751 CTGCTGGAGGGAGCCCGCTCCACACACACATCGCCCTTAGAGGGCGAGCCAT 810
Qy 1234 CTGGAGCAAGAGAGGATAAAGAGAGACACACCTCTTAGGGTCCCGAGGGGGCGG 1293
Db 811 CTGGAGCAAGAGAGGATAAAGAGAGACACACCTCTTAGGGTCCCGAGGGGGCGG 870
Qy 1294 GCGTGGGGTGGTAGGTGAAGAGCAG 1321
Db 871 GCGTGGGGTGGTAGGTGAAGAGCAG 898

RESULT 2

US-09-769-482-2

; Sequence 2, Application US/09769482

; Patent No. 6586130

; GENERAL INFORMATION:

; APPLICANT: SRIVASTAVA, SHIV

; APPLICANT: MOUL, JUDD W.

; APPLICANT: XU, LINDA L.

; APPLICANT: SEGAWA, TAKEHIKO

; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED

; FILE OF INVENTION: POYNUELOTIDE ARRAY

; FILE REFERENCE: 04995.0057-00000

; CURRENT APPLICATION NUMBER: US/09/769,482

; CURRENT FILING DATE: 2001-01-26

; PRIOR APPLICATION NUMBER: 60/178,772

; PRIOR FILING DATE: 2000-01-28

; PRIOR APPLICATION NUMBER: 60/179,045

; PRIOR FILING DATE: 2000-01-31

; NUMBER OF SEQ ID NOS: 67

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 2

; LENGTH: 759

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-769-482-2

Query Match

Best Local Similarity 57.2%; Score 755.8; DB 4; Length 759;

Matches 757; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 518 ATCACGGAGCTGAGTTGTTTCAGATCATCATCATCTGTTGATGATGATGATGATG 577

Db 1 ATCGCGGAGCTGAGTTGTTTCAGATCATCATCATCTGTTGATGATGATGATGATG 60

Qy 578 GTGGTATCAGTGCCTGCTGAGCCACTAAGAGCTGTGTGCAAGCTTCTTCAATCAGCCG 637
Db 61 GTGGTATCAGTGCCTGCTGAGCCACTAAGAGCTGTGTGCAAGCTTCTTCAATCAGCCG 120
Qy 638 CACAGCCAGGGCGGAGGAGAGATGCCCTGTCTCAGAAAGGATGCCCTGTGGCCCTCG 697
Db 121 CACAGCCAGGGCGGAGGAGAGATGCCCTGTCTCAGAAAGGATGCCCTGTGGCCCTCG 180
Qy 698 GAGAGCACAGTGTTCAGCAACGGAATCCAGAGCCGAGGTCTAAGCCCGGCTTGGCC 757
Db 181 GAGAGCACAGTGTTCAGCAACGGAATCCAGAGCCGAGGTCTAAGCCCGGCTTGGCC 240
Qy 758 ACCGAGCGCTGCGCTGCGCCCTTGGCGCAGCGGGGCGCTTCCACCGCTTCCAGCCC 817
Db 241 ACCGAGCGCTGCGCTGCGCCCTTGGCGCAGCGGGGCGCTTCCACCGCTTCCAGCCC 300
Qy 818 ACCTATCCGTACCTGAGCAGCAGATCGACCTGCCACCCACCATCTCGTGTGACAGCG 877
Db 301 ACCTATCCGTACCTGAGCAGCAGATCGACCTGCCACCCACCATCTCGTGTGACAGCG 360
Qy 878 GAGAGCCCGCCCTTACAGGGCCCTTGCACCTCCAGCTTGGGACCCCGAGCAGCAG 937
Db 361 GAGAGCCCGCCCTTACAGGGCCCTTGCACCTCCAGCTTGGGACCCCGAGCAGCAG 420
Qy 938 CTGGAACCTGAACCGGAGTGGTGGCGCACCCCAACAGAACCATCTTTCGACAGTGAC 997
Db 421 CTGGAACCTGAACCGGAGTGGTGGCGCACCCCAACAGAACCATCTTTCGACAGTGAC 480
Qy 998 CTGATGATAGTCCAGGCTGGCGGCGCTTGGCCCGCCCGCCAGCAGTAACTCGGGCATCAG 1057
Db 481 CTGATGATAGTCCAGGCTGGCGGCGCTTGGCCCGCCCGCCAGCAGTAACTCGGGCATCAG 540
Qy 1058 GCCACGTGTACCGGAGCGCGCGCATGGAGGGCGCGCCGCCACCTACAGCAGGTC 1117
Db 541 GCCACGTGTACCGGAGCGCGCGCATGGAGGGCGCGCCGCCACCTACAGCAGGTC 600
Qy 1118 ATCGGCCACTTACCGGGGCTCTCTTCCAGACCAAGAGGACCCCTCTTAG 1276
Db 601 ATCGGCCACTTACCGGGGCTCTCTTCCAGACCAAGAGGACCCCTCTTAG 660
Qy 1178 CTGGAGGGAGCCCGGTCCACACACACATCGCCCTTAGAGCGCAGCATCTCG 1237
Db 661 CTGGAGGGAGCCCGGTCCACACACACATCGCCCTTAGAGCGCAGCATCTCG 720
Qy 1238 AGCAAGAGAGGATTAACAGAAAGGACCCCTCTTAG 1276
Db 721 AGCAAGAGAGGATTAACAGAAAGGACCCCTCTTAG 759

RESULT 3

US-09-091-952A-7

; Sequence 7, Application US/09091952A

; Patent No. 6458532

; GENERAL INFORMATION:

; APPLICANT: Detera-Wadleigh, Sevilla D.

; Gershon, Elliot S.

; Badner, Judith A.

; Goldin, Lynn R.

; Berrettini, Wade H.

; Yoshikawa, Takeo

; Sanders, Alan R.

; Esterling, Lisa E.

; TITLE OF INVENTION: Chromosomal Markers and Diagnostic

; Tests for Manic-Depressive Illness

; NUMBER OF SEQUENCES: 197

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, Eighth Floor

; CITY: San Francisco

; STATE: CA

; COUNTRY: USA

; ZIP: 94111-3834

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/091,952A
FILING DATE: 19-Apr-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/029,278
FILING DATE: 28-OCT-1996
APPLICATION NUMBER: PCT/US97/19381
FILING DATE: 28-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Timothy L.
REGISTRATION NUMBER: 35,367
REFERENCE/DOCKET NUMBER: 015280-297100US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 921 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1...921
OTHER INFORMATION: Clone 22 coding region
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-091-952A-7

Query Match      26.7%; Score 352.2; DB 4; Length 921;
Best Local Similarity 73.1%; Pred. No. 4e-58;
Matches 482; Conservative 0; Mismatches 168; Indels 9; Gaps 2;

QY 506 CAGAGCATGAGATCAGGAGCTGGAGTTGTTCAGATCATCATCATCGTGGTGGTGATG 565
DB 166 CCGGGCATCTCACTCGGAGCTGGAGTTGCGCCAAATCATCATCATCGTGGTGGTGC 225
QY 566 ATGGTATGATGTTGGTGTATCAGTGCCTCTCAGGCACTACAAAGTGTCTGACGGTCC 625
DB 226 ACGGTATGTTGGTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 285
QY 626 TTCTATCAGCGGCACACCGGCGGCGGAGGAGAGATGCTGCTGCTGCTGCTGCTGCTG 685
DB 286 TTCTATCAGCGGCACACCGGCGGCGGAGGAGAGATGCTGCTGCTGCTGCTGCTGCTG 345
QY 686 CTGTGGCCCTCGGAGACACAGTGTCTAGGCAACGGAATCCAGAGCGCGAGGTCTACGCC 745
DB 346 CTGTGGCCCTCGGAGACACAGTGTCTAGGCAACGGAATCCAGAGCGCGAGGTCTACGCC 399
QY 746 CCGCTCGGCGCCACCGCTGCGTGGCTGCGTGGCTGCGTGGCTGCGTGGCTGCGTGGCT 805
DB 400 CCGCTCGGCGCCACCGCTGCGTGGCTGCGTGGCTGCGTGGCTGCGTGGCTGCGTGGCT 459
QY 806 CGCTTCAGCGCCACCTATCGGTACTCTGACGACGAGATCGACTGCGCCACCCACCATCTCG 865
DB 460 CGCTTCAGCGCCACCTATCGGTACTCTGACGACGAGATCGACTGCGCCACCCACCATCT 519
QY 866 CTGTTCAGCGGAGGAGGCGCCACCTATCGAGGCGCCCTGACCCCTCGAGTTCGGGAC 925
DB 520 CTGTTCAGCGGAGGAGGCGCCACCTATCGAGGCGCCCTGACCCCTCGAGTTCGGGAC 579
QY 926 CCGGAGCAGGAGTGGAGTGAACCGGAGTGCCTGCGCGCACCCCGGAGGAGGAGGAGGAG 985
DB 580 CCGGAGCAGGAGTGGAGTGAACCGGAGTGCCTGCGCGCACCCCGGAGGAGGAGGAGGAG 639
QY 986 TTCGACAGTGCACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1042
```

```
DB 640 TTTCACAGTGAATTAATAGACATTGCTATGATATAGCGGGGTCCATGCCACCCAGCAGC 699
QY 1043 AACTCGGGCATCAGCGCACCTGCTAGCGAGCGCGGGCGGCGCATGAGGGGCGCGCGCCC 1102
DB 700 AACTCGGGCATCAGTGCAGACACCTGCAGCAGTAACGAGGAGTGGAGGGGCCACCCCCC 759
QY 1103 ACCTACAGCAGGATCAGCGCCACTAGCGGGGTCTCTCTTCCAGCAGCAGCAGCAGCAG 1161
DB 760 ACATACAGCAGGATGATGGGCCACCCACCGGCGGCTCTTTCTCTCCATCACCAGCGCAG 818

RESULT 4
US-09-091-952A-6
; Sequence 6, Application US/09091952A
; Patent No. 6458532
; GENERAL INFORMATION:
; APPLICANT: Detera-Wadleigh, Sevilla D.
; Gershon, Elliot S.
; Badner, Judith A.
; Goldin, Lynn R.
; Berrettini, Wade H.
; Yoshikawa, Takeo
; Sanders, Alan R.
; Esterling, Lisa E.
; TITLE OF INVENTION: Chromosomal Markers and Diagnostic
; Tests for Manic-Depressive Illness
; NUMBER OF SEQUENCES: 197
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/091,952A
; FILING DATE: 19-Apr-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/029,278
; FILING DATE: 28-OCT-1996
; APPLICATION NUMBER: PCT/US97/19381
; FILING DATE: 28-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Timothy L.
; REGISTRATION NUMBER: 35,367
; REFERENCE/DOCKET NUMBER: 015280-297100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8065 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: -
; LOCATION: 1...8065
; OTHER INFORMATION: Clone 22
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 116...1036
; OTHER INFORMATION: Clone 22 coding region
; FEATURE:
```



```

QY 866 CTGTACAGCGGAGGAGCCACCTACAGGCGCCCTGACACCTCAGCTTCGGGAC 925
Db 466 CTGTCCGACGCTGAGAGACCACTCTTACAGGCGCCCTGACACCTCAGCTTCGGGAC 525
QY 926 CCGGAGCAGCAGCTGGAACCGGAGATCGGTGCGGCACACCCCAACAGAACATC 985
Db 526 CCTGAACAGCAGATGGAACCAACCGAGATCGGTGAGGCGCCCAACCAACCAACATA 585
QY 986 TTGACAGTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1042
Db 586 TTTGACAGTATTTAATAGACATTTGATATAGAGGAGTATGATGATGATGATGATGAT 645
QY 1043 AACTCGGCGATCAGCGCCAGCTGCTACGCGAGCGCGCGCGCATGAGGCGCGCGCGCC 1102
Db 646 AACTCGGCGATCAGTCAAGCAGCTGCGAGCAGTAAACGGAGAGTGGAGGGGCGCACCC 705
QY 1103 ACCTACAGCAGGTCATCGGCGCATCCCGGGGTCTCTTCCAGCAGCAGCAGAGAG 1161
Db 706 ACATACAGCAGGATGATGGGCGCCACCCAGGCGCTCTTTCTTCATCACCAGCGAG 764

```

RESULT 6

```

US-09-621-976-2162
; Sequence 2162, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2162
; LENGTH: 391
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 17..223
US-09-621-976-2162

```

```

Query Match 6.8%; Score 90.4; DB 4; Length 391;
Best Local Similarity 75.7%; Pred. No. 8.6e-09;
Matches 112; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 1019 GGCGGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1078
Db 26 GGGGGTCCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 85

QY 1079 GGGCGCATGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1138
Db 86 GGGAGGATGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 145

QY 1139 TCCTTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1166
Db 146 TCTTCTCCATCACCAGCGCAGACGC 173

```

RESULT 7

```

US-09-616-289-48/c
; Sequence 48, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING

```

```

; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 2561
; TYPE: DNA
; ORGANISM: Oryctolagus cuniculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (246)...(1895)
US-09-616-289-48

```

```

Query Match 5.7%; Score 75.4; DB 4; Length 2561;
Best Local Similarity 53.0%; Pred. No. 7.6e-06;
Matches 206; Conservative 0; Mismatches 181; Indels 2; Gaps 2;

QY 56 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 115
Db 875 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 816
QY 116 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 175
Db 815 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 756
QY 176 TGCCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 235
Db 755 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 697
QY 236 GCGACTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 295
Db 696 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 637
QY 296 CGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 355
Db 636 CGGCGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 578
QY 356 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 415
Db 577 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 518
QY 416 CACCGCTTGTATGGGGGTCAACAGCACCGC 444
Db 517 CTCCCGTTGTAGTACCGCGGAGCAGCGC 489

```

RESULT 8

```

US-09-616-289-48
; Sequence 48, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02

```

```

; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 2561
; TYPE: DNA
; ORGANISM: Oryctolagus cuniculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (246)...(1895)
US-09-616-289-48

Query Match          5.5%; Score 72.2; DB 4; Length 2561;
Best Local Similarity 51.1%; Pred. No. 3.1e-05;
Matches 170; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

QY 133 CGCCCGCTGCGAGCCCATTTCCGGACGCCACCCGCGGGGCACTGCCGAGCCCGCCGGG 192
Db 528 CGCAACGCGCGCGCGTCCAGCGCCCGCGCGCGAGCCACCCCGCGCGCCCGCGCGCG 587
QY 193 CTGCGAGGGAGGCGCGGGGGCGCGAGCGGAGCGCGTCCCGCGCACTGAGCCCGCGG 252
Db 588 GCCCGCGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 647
QY 253 CGCCCGCGGAATTGGCGGCGACCGCGAGCCCGCGAGCGCGGCGCGCTCCCGCGCGC 312
Db 648 CGCGCGCGCGCGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 707
QY 313 GGGCTCTGTCATGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 372
Db 708 GCGCTGCGCGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 767
QY 373 CGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 432
Db 768 CGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 827
QY 433 CAACAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 465
Db 828 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 860

RESULT 9
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (246)...(1895)
US-09-103-840A-2

Query Match          5.3%; Score 70.4; DB 3; Length 4403765;
Best Local Similarity 49.8%; Pred. No. 0.00039;
Matches 229; Conservative 0; Mismatches 227; Indels 4; Gaps 2;

QY 2 GACCGCGGTCTCGAGCGGAACCCGATCTCTTGGACTTGAATGAGGAGGAGCGCGC 61
Db 3942788 GCCGCGCGCAACCGCGCGCACCGCGCGCGAGCGCGAGTGTGCGCGCGCGCGCG 3942847
QY 62 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 121
Db 3942848 ATCGTGGCGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3942907
QY 122 GCGGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180

Query Match          5.2%; Score 68.8; DB 3; Length 441529;
Best Local Similarity 49.8%; Pred. No. 0.00039;
Matches 229; Conservative 0; Mismatches 227; Indels 4; Gaps 2;

QY 2 GACCGCGGTCTCGAGCGGAACCCGATCTCTTGGACTTGAATGAGGAGGAGCGCGC 61
Db 3942788 GCCGCGCGCAACCGCGCGCACCGCGCGCGAGCGCGAGTGTGCGCGCGCGCGCG 3942847
QY 62 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 121
Db 3942848 ATCGTGGCGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3942907
QY 122 GCGGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
```

Db	3942908	GGCGGCCCGCGCAGCGGCCCAACACAGTCCCGCGGCCAACGGCGGGCAAGAGGT	3942967
QY	181	-ACGCCCCCGGGACTGCCAGGGAGGCCCGGGGGGCGACGCGAGCGCGGTCCCGCGCA	239
Db	3942968	CAAGGCGCAGCGGTGTGTCGCTGGGGCGCCCGGGCTCGCGCGCCGCTGCGCGCT	3943027
QY	240	CTGAGCCCCGCGGCCCGCCCGGGAACTTGGGGGAGACCCGAGCCCGCGAGCCCGGGGGCGG	299
Db	3943028	AACGGACACCGCTGGCAACGCGCGCCAGGCGGTGTCGGCGGCAACCGCGCGCCCGCGCA	3943087
QY	300	CTTCCCGCCGCGCGGCTCTCTGATGCGGGGCCCCAGTCTCGGGCGCCCGCGGAGCCC	359
Db	3943088	GCCTCTTAGTACCAACGCGCGAGCGGGCGCGCCGG--CGGCACCGAGGCGACGGCG	3943145
QY	360	CCCCGGCGCCCCCGAGCCCCCGCGCCCGCGCGCGCGCGCGCGCTGCATGCACC	419
Db	3943146	GCAGCGCGCGCGCGCGCGCACTCGAGGCCGCCCGCGGCACCGCGGGCGGCGCGACG	3943205
QY	420	GCTTGATGGGGGTAAACAGACCGCGCCCGCCCGCCCGG	459
Db	3943206	GCGGCAAGATGGCCAGGGCGGGCGCCCGCGCGCGGTGCGG	3943245

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

? NAME/KEY: CDS
? LOCATION: 31329..36071
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 36155..41830
? -ID: US-08-804-227C-7

Query Match      5.1%; Score 68; DB 2; Length 44377;
Best Local Similarity 49.2%; Pred. No. 0.00029;
Matches 234; Conservative 0; Mismatches 240; Indels 2; Gaps 2;

QY 5 CGCGGCTCTCGGAGCGAAACCCGATCTCTCTGGACTTGAATGAGGAGGAGGAGGCGGCGGC 64
Db 19499 CTCGGGTGGGAACTCTCCGCGAGCGCGCGCAGCGCTCCGCGCGCGCTCGGCGGGC 19440
QY 65 GCGCGGCGCGCGGAGAGCGCTCGGCTGGGGAAGCTAGCGGCAAGAGCTAGCCCCGGCG 124
Db 19439 GGGGGCAGGCGCACGGCACCGCTCTCTCTCCCGGTGCCCCCGCGCCCGCGCGTGT 19380
QY 125 GCAAGCGCGCGCCCGCTGCGACGCCATTTTCGGAGCGCACCCCGCGGGCACTGCCGACGC 184
Db 19379 GCGGGCGTCCGCGGTAGCCCGGCCAGGTACGAGACCTTCGGGAGAGCGGTGGGCACC 19320
QY 185 CCCC GGCGGTGTCGAGGGGAGCGCGGGGGGCGCAGCGGAGCGCGTCCGCGCACTGAG 244
Db 19319 GGCGCGCGCGCGCAGGTGCGAGCGCACGGGCGAGCAGGGCCGCGTCTGGGCGCGGT-CAGCG 19261
QY 245 CCGCGGCGCGCCCGGGAACTTGGGGGCGACCCCGAGCCCGCGAGCGCGGGGCGCGCTCC 304
Db 19260 CCGGTCGAGCAGGTTCAGTGGCGCGCGGGGTGAGCGGGGCGAGCCCGAGGCGCTGA 19201
QY 305 CCGCGCGCGCGCTCTCTGATGCGGGGGCCCCAGGTCTCGGGGCGCGCGGAGCGCCCCCCC 364
Db 19200 CCGCGCGCGCGTCCGCTCCGTCCAGTCTGGTGGGAGACCCACCCCGGACCCACGAGCCCC 19141
QY 365 GCGCGCCCCCGAGACCCCCCGCGCGCCGCGCGCGCGCGCGTCTCATGSCACCGGTG 424
Db 19140 AGCCGAGCGACAGCGGGGTGTCCCGCGCGCGCACCCGCGGGCGGCCAGCGCTC-CATG 19082
QY 425 ATGGGGGTCAAAGCACCCCGCGCGCGCGCGCGCGCGCGCGCGCAATGTCTCTGCAAC 480
Db 19081 AAGGGGTTGGCGCGCGCGGTACCCGCGCTGCCCCCGGTGCGCGCGCGCGCGCGAGC 19026

```

RESULT 12
US-08-804-198-1/c
; Sequence 1, Application US/08804198
; Patent No. 5945320
; GENERAL INFORMATION:
; APPLICANT: BURGESS, Stanley G.
; APPLICANT: KUHSTOSS, Stuart A.
; APPLICANT: Rao, Nagaraja R.
; APPLICANT: Richardson, Mark A.
; APPLICANT: Rostock, Paul R., Jr.
; TITLE OF INVENTION: PLATENOLIDE SYNTHASE GENE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PAUL R. CANTRELL 1138
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,198
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

NAME: CANTRELL, PAUL R.
REGISTRATION NUMBER: 36,470
REFERENCE/DOCKET NUMBER: P9113
TELEPHONE: 317-276-3885
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 44377 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 350..14002
FEATURE:
NAME/KEY: CDS
LOCATION: 14046..20036
FEATURE:
NAME/KEY: CDS
LOCATION: 20110..31284
FEATURE:
NAME/KEY: CDS
LOCATION: 31329..36071
FEATURE:
NAME/KEY: CDS
LOCATION: 36155..41830
US-08-804-198-1

Query Match 5.1%; Score 68; DB 2; Length 44377;
Best Local Similarity 49.2%; Pred. No. 0.00029;
Matches 234; Conservative 0; Mismatches 240; Indels 2; Gaps 2;

QY 5 CCGCGTCTCGAGCGAAACCCGATCTCTTGACTTGAATGAGGAGGAGGCGGCGGC 64
Db 19499 CTCGGGTGCGGAACCTCCGCGAGCGCGCGCCAGCTCCGCGGGGTGCGCGGCGC 19440

QY 65 GCGCGGCGGCGGAGCGCTCGGTGGGAAAGTAGCGGAGAGGCTCAGCCCGCGCG 124
Db 19439 GGGGCGAGGCGCGACCGGACCGTCTCTGTCGCGGTGCGCGCGCGCGCGCGTCT 19380

QY 125 GCAGCG 184
Db 19379 GCGGCGGTGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 19320

QY 185 CCGCGGCGGTGCGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 244
Db 19319 GCG 19261

QY 245 CCG 304
Db 19260 CCGGTGCGAGCGGTCCAGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 19201

QY 305 CCG 364
Db 19200 CCG 19141

QY 365 GCG 424
Db 19140 AGCCGAGCGAGCGCGGCGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 19082

QY 425 ATGGGGGTCAACAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
Db 19081 AAGGGGTGGCGCGCGCGTACCGCGCGCGTCCCGCGCGCGTCCCGCGCGCGG 19026

RESULT 13

US-08-690-473-1/c
Sequence 1, Application US/08690473
Patent No. 5876923
GENERAL INFORMATION:
APPLICANT: Leopardi, Rosario
APPLICANT: Roizman, Bernard

TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 AS AN
TITLE OF INVENTION: INHIBITOR OF APOPTOSIS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/690,473
FILING DATE: 26-JUL-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: ARCD:239
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4257 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-690-473-1

Query Match 5.1%; Score 67.6; DB 2; Length 4257;
Best Local Similarity 48.8%; Pred. No. 0.00024;
Matches 239; Conservative 0; Mismatches 249; Indels 2; Gaps 2;

QY 40 TGAATGAGGAGGAGGAGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 99
Db 2828 TGCTGCG 2769

QY 100 TAGCGGAGAGGCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 159
Db 2768 GCGCGGAGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2709

QY 160 CGCCACCG 219
Db 2708 GCGTGGCG 2649

QY 220 GCGGAGCGCGGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 278
Db 2648 GGGCGCGCGCGGTCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2589

QY 279 AGCCCG 338
Db 2588 GGGCG 2529

QY 339 TCGG-GGCG 397
Db 2528 GGGCG 2469

QY 398 GCG 457
Db 2468 GCG 2409

QY 458 GGGCAGCG 517
Db 2408 GCGGAGCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2349

QY 518 ATCAGCGGAGC 527
Db 2348 CTCAGCGGCGC 2339

Search completed: May 12, 2004, 05:21:48
Job time : 152.422 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2004, 03:27:13 ; Search time 639,311 Seconds
(without alignments)
9360.266 Million cell updates/sec

Title: US-09-934-249-1

Perfect score: 1321

Sequence: 1 CGACCGGGTCTCGGAGCGA.....CTGCGTAGGTGAAAGGCGAG 1321

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2941586 seqs, 2264995651 residues

Total number of hits satisfying chosen parameters: 5883172

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1321	100.0	1321	9	US-09-934-249-1
2	1229	93.0	4839	15	US-10-241-220-119
3	1229	93.0	4839	15	US-10-269-909-84
4	1229	93.0	4839	15	US-10-269-909-85
5	864	65.4	864	16	US-10-295-027-127
6	861	65.2	861	9	US-09-934-249-3
7	803.2	60.8	969	10	US-09-796-753-55
8	802.2	60.7	1141	15	US-10-205-823-412
9	802.2	60.7	1141	15	US-10-301-822-208
10	802.2	60.7	1850	15	US-10-241-220-44
11	802.2	60.7	4527	10	US-09-821-812-2
12	801.6	60.7	1140	15	US-10-390-045-1
13	790.4	59.8	1066	14	US-10-098-841-71
14	767.4	58.1	806	15	US-10-241-220-45

15	755.8	57.2	759	15	US-10-390-045-2	Sequence 12, Appli
16	616.4	46.7	878	9	US-09-934-249-12	Sequence 12, Appli
17	583.2	44.6	1583	15	US-10-000-256A-32	Sequence 32, Appl
18	493.6	37.4	693	9	US-09-934-249-14	Sequence 14, Appl
19	466.2	35.3	1713	10	US-09-796-753-57	Sequence 57, Appl
20	352.2	26.7	921	15	US-10-251-598-7	Sequence 0, Appli
21	352.2	26.7	8018	13	US-10-087-192-1667	Sequence 1667, Ap
22	352.2	26.7	8065	15	US-10-251-598-6	Sequence 0, Appli
23	350	26.5	920	13	US-09-823-245A-20	Sequence 20, Appl
24	350	26.5	2170	16	US-10-094-749-137	Sequence 197, App
25	349	26.4	8093	9	US-09-934-249-16	Sequence 16, Appl
26	349	26.4	8093	13	US-10-343-887-1023	Sequence 1023, Ap
27	349	26.4	8093	13	US-10-172-118-1023	Sequence 1023, Ap
28	341.8	25.9	475	9	US-09-934-249-15	Sequence 15, Appl
29	284.2	21.5	867	15	US-10-251-598-8	Sequence 0, Appli
30	250.2	18.9	285020	13	US-10-087-192-1666	Sequence 1666, Ap
31	182	13.8	368	9	US-09-783-590-346A	Sequence 346A, Ap
32	182	13.8	467	10	US-09-918-995-2074	Sequence 2074, Ap
33	100.2	7.6	522	13	US-10-240-425-166	Sequence 166, App
34	79.6	6.0	51657	16	US-10-057-475B-10475	Sequence 10475, A
35	79.6	6.0	51657	16	US-10-154-884B-10475	Sequence 10475, A
36	78	5.9	4020	10	US-09-796-679-5	Sequence 5, Appli
37	77.6	5.9	606	13	US-10-424-593-95310	Sequence 95310, A
38	75.4	5.7	2581	9	US-09-976-740-48	Sequence 48, Appl
39	75.4	5.7	2581	13	US-10-671-242-48	Sequence 48, Appl
40	75.4	5.7	2561	14	US-10-023-523-48	Sequence 48, Appl
41	75.4	5.7	2561	14	US-10-023-523-48	Sequence 48, Appl
42	75.4	5.7	2561	16	US-10-616-187-48	Sequence 48, Appl
43	72.2	5.5	2307	10	US-09-893-519A-87	Sequence 87, Appl
44	72.2	5.5	2561	9	US-09-976-740-48	Sequence 48, Appl
45	72.2	5.5	2561	13	US-10-671-242-48	Sequence 48, Appl

ALIGNMENTS

RESULT 1

US-09-934-249-1
; Sequence 1, Application US/09934249
; Patent No. US20020115081A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Richard T.
; APPLICANT: Landschulz, Katherine T.
; APPLICANT: Turi, Thomas G.
; APPLICANT: Thompson, John F.
; APPLICANT: Kennedy, Scott P.
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF
; FILE REFERENCE: P0738/7001/ERP/KA
; CURRENT APPLICATION NUMBER: US/09/934,249
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/227,159
; PRIOR FILING DATE: 2000-08-22
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1321
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (413) ... (1273)
US-09-934-249-1

Query Match 100.0%; Score 1321; DB 9; Length 1321;
Best Local Similarity 100.0%; Pred. No. 5.3e-293;
Matches 1321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACCGGGTCTCGGAGCGAACCAGATCTCTTGGACTTGAATGAGGAGGAGCGG 60
Db 1 CGACCGGGTCTCGGAGCGAACCAGATCTCTTGGACTTGAATGAGGAGGAGCGG 60
QY 61 CGGCGGGCGGCGGCGGCGGCTCGGCTGGGGAAGACTAGCGGAGAGCTCAGCCCC 120

421 TGAGATCAGCGAGCTGGAGTTGTTTCAGATCATCATCATCTGTTGGTGTGATGGTGA 480
573 TGGTGTGTGTATCATCGTGCCTGCTTACGACCTATCAAGCTGTGACAGTCTTTCATCA 532
481 TGGTGTGTGTATCATCGTGCCTGCTTACGACCTATCAAGCTGTGACAGTCTTTCATCA 540
633 GCCGGCACAGCCAGCGGCGGAGAGAGATGCTGCTTCCCTCAGAAAGATGCTGTGGC 692
541 GCCGGCACAGCCAGCGGCGGAGAGAGATGCTGCTTCCCTCAGAAAGATGCTGTGGC 600
593 CTTCCGAGAGACAGTGTTCAGGCAACGGAAATCCAGAGCGGACGAGTCTACGCGCCCTC 752
601 CTTCCGAGAGACAGTGTTCAGGCAACGGAAATCCAGAGCGGACGAGTCTACGCGCCCTC 660
753 GGCCACACAGCGGCGGAGAGAGATGCTGCTTCCCTCAGAAAGATGCTGTGGC 812
661 GGCCACACAGCGGCGGAGAGAGATGCTGCTTCCCTCAGAAAGATGCTGTGGC 720
813 AGCCACCTATCTGCTACCTGACGACGAGATGCTGCTTCCCTCAGAAAGATGCTGTGGC 872
721 AGCCACCTATCTGCTACCTGACGACGAGATGCTGCTTCCCTCAGAAAGATGCTGTGGC 780
873 AGCCACCTATCTGCTACCTGACGACGAGATGCTGCTTCCCTCAGAAAGATGCTGTGGC 932
781 AGCCACCTATCTGCTACCTGACGACGAGATGCTGCTTCCCTCAGAAAGATGCTGTGGC 840
933 AGCAGCTGGAATGAAACCGGAGTCTGCTGCGGACCCCTCAGAAAGATGCTGTGGC 992
841 AGCAGCTGGAATGAAACCGGAGTCTGCTGCGGACCCCTCAGAAAGATGCTGTGGC 900
993 GTGACCTGATGATAGTGGAGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1052
901 GTGACCTGATGATAGTGGAGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
1053 TCAGCGCCAGCTGCTACGCGAGCGGCGGCGCATGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1112
961 TCAGCGCCAGCTGCTACGCGAGCGGCGGCGCATGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
1113 AGCTATCGGCGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1172
1021 AGCTATCGGCGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
1173 CTTTGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1232
1081 CTTTGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
1233 TCTGGAGCAAGAGAGAGGATTAACAGAAAGGACACCTCTCTTAGGGTCTCCAGGGGGGCGC 1292
1141 TCTGGAGCAAGAGAGAGGATTAACAGAAAGGACACCTCTCTTAGGGTCTCCAGGGGGGCGC 1200
1293 GGGCTGGGGCTGCTAGGTGAAAGGCGAG 1321
1201 GGGCTGGGGCTGCTAGGTGAAAGGCGAG 1229

RESULT 3

US-10-269-909-84
; Sequence 84, Application US/10269909
; Publication No. US20030180747A1
; GENERAL INFORMATION:
; APPLICANT: HRUBAN, RALPH H.
; APPLICANT: ARGANI, PEDRAM
; APPLICANT: IACOBUZIO-DONAHUE, CHRISTINE
; APPLICANT: MAITRA, ANIRBAN
; TITLE OF INVENTION: PANCREATIC CANCER DIAGNOSIS AND THERAPIES
; FILE REFERENCE: 58303(71699)
; CURRENT APPLICATION NUMBER: US/10269,909
; CURRENT FILING DATE: 2003-10-11
; PRIOR APPLICATION NUMBER: 60/328,609
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 60/332,754
; PRIOR FILING DATE: 2001-11-19

; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 84
; LENGTH: 4839
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-269-909-84

Query Match 93.0%; Score 1229; DB 15; Length 4839;
Best Local Similarity 100.0%; Pred. No. 6.3e-272;
Matches 1229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 GGAAGCTAGCGGAGAGGCTACGCCCCGGGGGAGCGCGGCCCCCTGCGACGCCATT 152
DB 1 GGAAGCTAGCGGAGAGGCTACGCCCCGGGGGAGCGCGGCCCCCTGCGACGCCATT 60
QY 153 TTCCGGAGCGCCACCCCGGGGCACTGCCGACGCGCCCCGGGGTGCAGAGGGAGGCGCGGG 212
DB 51 TTCCGGAGCGCCACCCCGGGGCACTGCCGAGCGCCCCGGGGTGCAGAGGGAGGCGCGGG 120
QY 213 GGGCGAGCGAGGCGCGGCTCCCGGCACTGAGCCCCCGGGGCGCCCGGGAACTTTGGCGGC 272
DB 121 GGGCGAGCGAGGCGCGGCTCCCGGCACTGAGCCCCCGGGGCGCCCGGGAACTTTGGCGGC 180
QY 273 GACCGGAGCGCGGAGCGCGGCGCGCTCCCGGCGCGCGCGCTCTCTGATGCGGGGC 332
DB 181 GACCGGAGCGCGGAGCGCGGCGCGCTCCCGGCGCGCGCGCTCTCTGATGCGGGGC 240
QY 333 CCCAGCTCCGGGCGCGCGGAGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 392
DB 241 CCCAGCTCCGGGCGCGCGGAGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 393 GCGCGCGCGCGCGCGCGCTTCCATGACCGCTTGTAGGGGGTCAACAGACCGCGCGCGCGCG 452
DB 301 GCGCGCGCGCGCGCGCGCTTCCATGACCGCTTGTAGGGGGTCAACAGACCGCGCGCGCGCG 360
QY 453 CCGCGCGCGCGCGCGCGCTTCTCTGCGAGTGCACCTGCAACCGCTTGTTCAGAGCA 512
DB 361 CCGCGCGCGCGCGCGCGCTTCTCTGCGAGTGCACCTGCAACCGCTTGTTCAGAGCA 420
QY 513 TGGAGATCAGCGAGCTGAGGTTTGTTCAGATCATCATCATCGTGGTGTGATGATGGA 572
DB 421 TGGAGATCAGCGAGCTGAGGTTTGTTCAGATCATCATCATCGTGGTGTGATGATGGA 480
QY 573 TGGTGGTGGTATCAGCTGCTGTGAGCCACTAAGCTGTCTGCAAGTCTTTCATCA 632
DB 481 TGGTGGTGGTATCAGCTGCTGTGAGCCACTAAGCTGTCTGCAAGTCTTTCATCA 540
QY 633 GCGGCA CAGCCAGCGGCGGAGGAGGAGGATGCGCTTCTCAGAGGATGCGCTGGC 692
DB 541 GCGGCA CAGCCAGCGGCGGAGGAGGAGGATGCGCTTCTCAGAGGATGCGCTGGC 600
QY 693 CCGCGAGGACAGTGTGAGGCAACGGAATCCAGAGCGCGAGTCTACCGCGCGCGCTC 752
DB 601 CCGCGAGGACAGTGTGAGGCAACGGAATCCAGAGCGCGAGTCTACCGCGCGCGCTC 660
QY 753 GCGCCACCGAGCGGCTGCGCGCTTTCGCGCGAGCGGAGGCGCTTTCAGCGCTTTC 812
DB 661 GCGCCACCGAGCGGCTGCGCGCTTTCGCGCGAGCGGAGGCGCTTTCAGCGCTTTC 720
QY 813 AGCCACCTATCTGCTACCTGACGACGAGATGCTGCTTCCCTCAGAAAGATGCTGTGGC 872
DB 721 AGCCACCTATCTGCTACCTGACGACGAGATGCTGCTTCCCTCAGAAAGATGCTGTGGC 780
QY 873 AGCGGAGGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 932
DB 781 AGCGGAGGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
QY 933 AGCAGCTGGAATGAAACCGGAGTCTGCTGCGGACCCCTCAGAAAGATGCTGTGGC 992
DB 841 AGCAGCTGGAATGAAACCGGAGTCTGCTGCGGACCCCTCAGAAAGATGCTGTGGC 900
QY 993 GTGACCTGATGATAGTGGAGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1052

Db 901 GTGACCTGATGATAGTCCAGAGCTGGGCGGCCCTTGCCTCCCGCCAGCAGTAATCTCGGGCA 960
Qy 1053 TCAGCGCCACGTGTCTAGCGCAGCGCGCGGCATAGAGGGGCGCCGCCACCTACAGCG 1112
Db 961 TCAGCGCCACGTGTCTAGCGCAGCGCGCGGCATAGAGGGGCGCCGCCACCTACAGCG 1020
Qy 1113 AGGTATCGGCGCACTACCGGGGTCCTCTTCCAGCACCAGCAGCAGTGGGCGGCCCT 1172
Db 1021 AGGTATCGGCGCACTACCGGGGTCCTCTTCCAGCACCAGCAGCAGTGGGCGGCCCT 1080
Qy 1173 CTTGCTGAGGGGACCGGCTCCACACACACACATCGCGCCCTCTAGAGCGCAGCCA 1232
Db 1081 CTTGCTGAGGGGACCGGCTCCACACACACACATCGCGCCCTCTAGAGCGCAGCCA 1140
Qy 1233 TCTGAGCAAGAGAGAGATAAACAGAAAGACACCTCTCTAGAGTCCCGAGGGGGCC 1292
Db 1141 TCTGAGCAAGAGAGAGATAAACAGAAAGACACCTCTCTAGAGTCCCGAGGGGGCC 1200
Qy 1293 GGGCTGGGGCTCGGTAGGTGAAAGGCAG 1321
Db 1201 GGGCTGGGGCTCGGTAGGTGAAAGGCAG 1229

RESULT 4
US-10-269-909-85
; Sequence 85, Application US/10269909
; Publication No. US20030180747A1
; GENERAL INFORMATION:
; APPLICANT: HRUBAN, RALPH H.
; APPLICANT: ARGANI, PEDRAM
; APPLICANT: IACOBUTIO-DONAHUE, CHRISTINE
; APPLICANT: MAITRA, ANIRBAN
; TITLE OF INVENTION: PANCREATIC CANCER DIAGNOSIS AND THERAPIES
; FILE REFERENCE: 58303(71699)
; CURRENT APPLICATION NUMBER: US/10/269,909
; CURRENT FILING DATE: 2003-10-11
; PRIOR APPLICATION NUMBER: 60/328,609
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 60/332,754
; PRIOR FILING DATE: 2001-11-19
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 85
; LENGTH: 4839
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-269-909-85

Query Match 93.0%; Score 1229; DB 15; Length 4839;
Best Local Similarity 100.0%; Pred. No. 6.3e-272;
Matches 1229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 93 GGAAGCTAGCGGCGAGGCTCAGCCCGCGGCGAGCGCGCGCCCGCTGCGAGCCATT 152
Db 1 GGAAGCTAGCGGCGAGGCTCAGCCCGCGGCGAGCGCGCGCCCGCTGCGAGCCATT 60
Qy 153 TTCGGACCCACCGCGGCGACTGCCGAGCCCGCGGCTGCCGAGGGAGGCGCGGG 212
Db 61 TTCGGACCCACCGCGGCGACTGCCGAGCCCGCGGCTGCCGAGGGAGGCGCGGG 120
Qy 213 GGGCGAGCGGAGCGGCTCCCGGCACTGAGCCCGCGGCGCCCGCGGAACTTGGCGGC 272
Db 121 GGGCGAGCGGAGCGGCTCCCGGCACTGAGCCCGCGGCGCCCGCGGAACTTGGCGGC 180
Qy 273 GACCCGAGCCCGCGGAGCGGCGCGCTCCCGCGGCGCGCTCTGATGATGGGGGC 332
Db 181 GACCCGAGCCCGCGGAGCGGCGCGCTCCCGCGGCGCGCTCTGATGATGGGGGC 240
Qy 333 CCCAGTCCGGCGCGCGCGGAGCGCCCGCGGCGCGCCCGCGGCGCGCCCGCGCGCC 392
Db 241 CCCAGTCCGGCGCGCGCGGAGCGCCCGCGGCGCGCCCGCGGCGCGCCCGCGCGCC 300

Qy 393 GCGCGCGCGCGCGCGCTCATGCACTGCTTGAATGGGGGTCAAACAGACCGCGCGCGCG 452
Db 301 GCGCGCGCGCGCGCGCTCATGCACTGCTTGAATGGGGGTCAAACAGACCGCGCGCGCG 360
Qy 453 CCGCGCGCGCGAGCCCAATGTCTCTGCACTGCAAGTGAATGCAAAAGCGTCTTTTTCAGAGCA 512
Db 361 CCGCGCGCGCGAGCCCAATGTCTCTGCACTGCAAGTGAATGCAAAAGCGTCTTTTTCAGAGCA 420
Qy 513 TGGAGTCAAGAGCTGGAGTTTGTTCAGATCATCATCTGCTGCTGCTGATGATGGA 572
Db 421 TGGAGTCAAGAGCTGGAGTTTGTTCAGATCATCATCTGCTGCTGCTGATGATGGA 480
Qy 573 TGGTGTGTGATCAGCTGCTGCTGAGCCTACAGCTGTCTGCAAGTCTCTTTCATCA 632
Db 481 TGGTGTGTGATCAGCTGCTGCTGAGCCTACAGCTGTCTGCAAGTCTCTTTCATCA 540
Qy 633 GCGCGCAGCCAGCGGCGGAGAGAGATGCCCTCTCTCAGAAAGTGCCTGTGGC 692
Db 541 GCGCGCAGCCAGCGGCGGAGAGAGATGCCCTCTCTCAGAAAGTGCCTGTGGC 600
Qy 693 CTTGAGAGCAGCAGTGTGAGGCAACGGAAATCCAGAGCGCGAGGTCTACGCGCGCGCTC 752
Db 601 CTTGAGAGCAGCAGTGTGAGGCAACGGAAATCCAGAGCGCGAGGTCTACGCGCGCGCTC 660
Qy 753 GCGCCACCGACCGCTGCGCGCTTGCCTTGCCTCAGCGGAGCGCTTCCACCGCTTCC 812
Db 661 GCGCCACCGACCGCTGCGCGCTTGCCTTGCCTCAGCGGAGCGCTTCCACCGCTTCC 720
Qy 813 AGCCCACTATCCGTACCTGCGAGCAGATGACCTGCCACCCACCATCTCGCTGTGAG 872
Db 721 AGCCCACTATCCGTACCTGCGAGCAGATGACCTGCCACCCACCATCTCGCTGTGAG 780
Qy 873 AGGGGAGGAGCGCCACCTTACAGGGCGCTTGCACCTCTCAGCTTGGGAGCGCGAGC 932
Db 781 AGGGGAGGAGCGCCACCTTACAGGGCGCTTGCACCTCTCAGCTTGGGAGCGCGAGC 840
Qy 933 AGCAGCTGAACTGAAACCGGGAGTGGTGGCGGCGACCCCAACAGAACCATCTTCGACA 992
Db 841 AGCAGCTGAACTGAAACCGGGAGTGGTGGCGGCGACCCCAACAGAACCATCTTCGACA 900
Qy 993 GTGACCTGATGATAGTCCAGGCTGGGGCGGCGCTTGCCTTGCCTCAGCTTGGGAGC 1052
Db 901 GTGACCTGATGATAGTCCAGGCTGGGGCGGCGCTTGCCTTGCCTCAGCTTGGGAGC 960
Qy 1053 TCAGCGCCACGTGTACGCGCGGCGCATGGAGGGCGCGCGCGCCACCTACCTACAGCG 1112
Db 961 TCAGCGCCACGTGTACGCGCGGCGCATGGAGGGCGCGCGCGCCACCTACCTACAGCG 1020
Qy 1113 AGTTCATCGGCCCACTACCGCGGGTCTCTTTCAGCACACAGCAGAGTGGGCGCGCT 1172
Db 1021 AGTTCATCGGCCCACTACCGCGGGTCTCTTTCAGCACACAGCAGAGTGGGCGCGCT 1080
Qy 1173 CTTGCTGAGGGGAGCGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCCA 1232
Db 1081 CTTGCTGAGGGGAGCGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCCA 1140
Qy 1233 TCTGGAGCAAAAGAGAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCGAGGGGGCC 1292
Db 1141 TCTGGAGCAAAAGAGAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCGAGGGGGCC 1200
Qy 1293 GGGCTGGGGCTCGGTAGGTGAAAGGCAG 1321
Db 1201 GGGCTGGGGCTCGGTAGGTGAAAGGCAG 1229

RESULT 5
US-10-295-027-127
; Sequence 127, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.


```
Db 242 GCCACCGAGCGCTGCGCTGCGCTTGGCCAGCGGAGCGTCCACCGCTTCCA 301
Qy 814 GCCACCTATCGTACTCGACGACGAGATCGACTGCGACCCACGATCTCGCTGAGA 873
Db 302 GCCACCTATCGTACTCGACGACGAGATCGACTGCGCGCCACCATCTGCTGAGA 361
Qy 874 CGGGGAGGAGCGCCCAACCTTACAGGGGCGGCTGCGACCTTCCAGCTTCCGAGCCCGGAGCA 933
Db 362 CGGGGAGGAGCGCCCAACCTTACAGGGGCGGCTGCGACCTTCCAGCTTCCGAGCCCGGAGCA 421
Qy 934 GCAGCTGGAAGTGAACCGGGAGTGGTGGCGGACCGGCGGACCGGCGGACCGGCGGAG 993
Db 422 GCAGCTGGAAGTGAACCGGGAGTGGTGGCGGACCGGCGGACCGGCGGACCGGCGGAG 481
Qy 994 TGACCTGATGGATAGTGGCGGAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAT 1053
Db 482 TGACCTGATGGATAGTGGCGGAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAT 541
Qy 1054 CAGCGCACGCTGTACCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAG 1113
Db 542 CAGCGCACGCTGTACCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAG 601
Qy 1114 GGTATCTGGGCACTACCGGGGTCCTCTTCCAGCACGAGGAGGAGGAGGAGGAGGAGGAG 1173
Db 502 GGTATCTGGGCACTACCGGGGTCCTCTTCCAGCACGAGGAGGAGGAGGAGGAGGAGGAG 561
Qy 1174 CTTGCTGGAGGAGCGCGGCTCCACACACACATCGCGCGGCTTAGAGAGCGGAGCCAT 1233
Db 662 CTTGCTGGAGGAGCGCGGCTCCACACACACATCGCGCGGCTTAGAGAGCGGAGCCAT 721
Qy 1234 CTTGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1293
Db 722 CTTGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 781
Qy 1294 GGCTGGGGTGGCTAGGTGAAAAGGCGAG 1321
Db 782 GGCTGGGGTGGCTAGGTGAAAAGGCGAG 809
```

RESULT 8

```
US-10-205-823-412
; Sequence 412, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Monsey, Angela M.
; APPLICANT: Glatz, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MEI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
```

RESULT 9

```
US-10-301-822-208
; Sequence 208, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 412
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-205-823-412
```

Query Match 60.7%; Score 802.2; DB 15; Length 1141;

Best Local Similarity 98.4%; Pred. No. 3e-174;

Matches 810; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```
Qy 499 TTTGTTCCAGAGCATGAGATCAGGAGCTGGAGTTTGTTCAGATCATCATCTCGTGGT 558
Db 77 TCTCTCGAAACAGGCAATGGCGGAGCTGGAGTTTGTTCAGATCATCATCTCGTGGT 136
Qy 559 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 618
Db 137 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 196
Qy 619 ACGTGCTTTCATCAGCGGACAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 678
Db 197 ACGTGCTTTCATCAGCGGACAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 256
Qy 679 AGGATGCTGTGGGCTTCGGAGAGCAGATGTCAAGCAACGGAATCCAGAGCGCAGGT 738
Db 257 AGGATGCTGTGGGCTTCGGAGAGCAGATGTCAAGCAACGGAATCCAGAGCGCAGGT 316
Qy 739 CTAGCGCGGCTTCGGGCGGACCGGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCG 798
Db 317 CTAGCGCGGCTTCGGGCGGACCGGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCG 376
Qy 799 CTTCCACCGCTTCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 858
Db 377 CTTCCACCGCTTCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 436
Qy 859 CATCTCGCTGTCAAGCGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 918
Db 437 CATCTCGCTGTCAAGCGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 496
Qy 919 TCGGAGCGGCGGAGCAGGCTGGAACCGGGAGTGGTGGCGGCGGCGGCGGCGGCGGCGGCG 978
Db 497 TCGGAGCGGCGGAGCAGGCTGGAACCGGGAGTGGTGGCGGCGGCGGCGGCGGCGGCGGCG 556
Qy 979 AACCATCTTCACAGTGAACCTGTGATGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1038
Db 557 AACCATCTTCACAGTGAACCTGTGATGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 616
Qy 1039 CAGTAACTCGGGCATCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1098
Db 617 CAGTAACTCGGGCATCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 676
Qy 1099 GCCCACCTTACAGGAGGTGATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1158
Db 677 GCCCACCTTACAGGAGGTGATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 736
Qy 1159 CAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1218
Db 737 CAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 796
Qy 1219 AGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1278
Db 797 AGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 856
Qy 1279 TCCCCAGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1321
Db 857 TCCCCAGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 899
```



```
Db 437 CATCTCGTGTACAGCGGGAGAGCCCCACCCCTACCAAGGCCCTGACCCCTCAGCT 496
QY 919 TCGGACCCCGAGCAGCAGCTGGAATCTAACCGGAGTCTGGTGGGCGACCCCGAAACAG 978
Db 497 TCGGACCCCGAGCAGCAGCTGGAATCTAACCGGAGTCTGGTGGGCGACCCCGAAACAG 556
QY 979 AACCATCTTCGACAGTGAACCTGATGATAGTGCAGGCTGGGCGGCCCTGCCCCCCCAG 1038
Db 557 AACCATCTTCGACAGTGAACCTGATGATAGTGCAGGCTGGGCGGCCCTGCCCCCCCAG 616
QY 1039 CAGTAACCTCGGAGCAGCAGCTGGAATCTAACCGGAGTCTGGTGGGCGACCCCGAAACAG 1098
Db 617 CAGTAACCTCGGAGCAGCAGCTGGAATCTAACCGGAGTCTGGTGGGCGACCCCGAAACAG 676
QY 1099 GCCCACTACAGCAGTGAACCTGATGATAGTGCAGGCTGGGCGGCCCTGCCCCCCCAG 1158
Db 677 GCCCACTACAGCAGTGAACCTGATGATAGTGCAGGCTGGGCGGCCCTGCCCCCCCAG 736
QY 1159 CAGTGGGCCGCCCTTCCTTGTGAGGGGACCCGGCTCCACCAACACATCGCGGCCCT 1218
Db 737 CAGTGGGCCGCCCTTCCTTGTGAGGGGACCCGGCTCCACCAACACATCGCGGCCCT 796
QY 1219 AGAGAGCGCAGCAGCTCTGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1278
Db 797 AGAGAGCGCAGCAGCTCTGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 856
QY 1279 TCCCCAGGGGGCGCGGCTGGGCTGCGTAGGTGAAAGGCAG 1321
Db 857 TCCCCAGGGGGCGCGGCTGGGCTGCGTAGGTGAAAGGCAG 899

RESULT 11
US-09-821-812-2
; Sequence 2, Application US/09821812
; Publication No. US20030166520A1
; GENERAL INFORMATION:
; APPLICANT: Lin, Biaoyang
; TITLE OF INVENTION: Androgen Regulated Prostate Specific
; FILE REFERENCE: P-IS 4373
; CURRENT APPLICATION NUMBER: US/09/821,812
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 4527
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (96)...(851)
US-09-821-812-2
Query Match 60.7%; Score 802.2; DB 10; Length 4527;
Best Local Similarity 98.4%; Pred. No. 3.2e-174;
Matches 810; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 499 TTGTTTCCAGCAGCATGAGATCAGGAGTGGAGTTGTTTCAGATCATCATCATGCTGT 558
Db 77 TCTCTTCGAAACACAGCAGTATGCGAGCTGGAGTTGTTTCAGATCATCATCATGCTGT 136
QY 559 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 618
Db 137 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 196
QY 619 ACGGTCCTTCATCAGCGGCACACCCAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 678
Db 197 ACGGTCCTTCATCAGCGGCACACCCAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 256
QY 679 AGGATGCTGTGGCCCTTCGAGAGCAGCAGTGTGAGGCAACGGAATCCAGAGCGCGAGGT 738
Db 257 AGGATGCTGTGGCCCTTCGAGAGCAGCAGTGTGAGGCAACGGAATCCAGAGCGCGAGGT 316
```

```
QY 739 CTACGCCCCGCTCGGCCCCACCGACCGCCTTGGCGTGGCGCCCTTCCGCCAGCGGGAGCG 798
Db 317 CTACGCCCCGCTCGGCCCCACCGACCGCCTTGGCGTGGCGCCCTTCCGCCAGCGGGAGCG 376
QY 799 CTTCCACCGCTTCAGCGCCACCTATCCGTAACCTGACGACGAGATCGACCTGCCACCCAC 858
Db 377 CTTCCACCGCTTCAGCGCCACCTATCCGTAACCTGACGACGAGATCGACCTGCCACCCAC 436
QY 859 CATCTCGCTGTACAGCGGGAGGAGCCGCCACCTACAGGGCCCTGACCCCTCCAGCT 918
Db 437 CATCTCGCTGTACAGCGGGAGGAGCCGCCACCTACAGGGCCCTGACCCCTCCAGCT 496
QY 919 TCGGAGACCCCGAGCAGCAGCTGGAATCTAACCGGAGTCTGGTGGGCGACCCCGAAACAG 978
Db 497 TCGGAGACCCCGAGCAGCAGCTGGAATCTAACCGGAGTCTGGTGGGCGACCCCGAAACAG 556
QY 979 AACCATCTTCGACAGTGAACCTGATGATAGTGCAGGCTGGGCGGCCCTGCCCCCCCAG 1038
Db 557 AACCATCTTCGACAGTGAACCTGATGATAGTGCAGGCTGGGCGGCCCTGCCCCCCCAG 616
QY 1039 CAGTAACCTCGGAGCAGCAGCTGGAATCTAACCGGAGTCTGGTGGGCGACCCCGAAACAG 1098
Db 617 CAGTAACCTCGGAGCAGCAGCTGGAATCTAACCGGAGTCTGGTGGGCGACCCCGAAACAG 676
QY 1099 GCCCACTACAGCAGTGAACCTGATGATAGTGCAGGCTGGGCGGCCCTGCCCCCCCAG 1158
Db 677 GCCCACTACAGCAGTGAACCTGATGATAGTGCAGGCTGGGCGGCCCTGCCCCCCCAG 736
QY 1159 CAGTGGGCCGCCCTTCCTTGTGAGGGGACCCGGCTCCACCAACACATCGCGGCCCT 1218
Db 737 CAGTGGGCCGCCCTTCCTTGTGAGGGGACCCGGCTCCACCAACACATCGCGGCCCT 796
QY 1219 AGAGAGCGCAGCAGCTCTGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1278
Db 797 AGAGAGCGCAGCAGCTCTGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 856
QY 1279 TCCCCAGGGGGCGCGGCTGGGCTGCGTAGGTGAAAGGCAG 1321
Db 857 TCCCCAGGGGGCGCGGCTGGGCTGCGTAGGTGAAAGGCAG 899

RESULT 12
US-10-390-045-1
; Sequence 1, Application US/10390045
; Publication No. US20030170713A1
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, SHIV
; APPLICANT: MOUL, JUDD W.
; APPLICANT: XU, LINDA L.
; APPLICANT: SEGAWA, TAKEHIKO
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED
; FILE REFERENCE: 04995.0057-00000
; CURRENT APPLICATION NUMBER: US/10/390,045
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US/09/769,482
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,045
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1140
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (95)...(850)
US-10-390-045-1
```


Db 653 CGTGTACGGCAGCGCGGGCGCATGGAGGGCGCGCGCCCGCCACCTACAGCGAGGTATCG 712
 QY 1122 GCCACTACCCGGGGTCTCTCTCCAGCACCAGCAGCAGTGGCGCGCCCTCTCTTGGTGG 1181
 Db 713 GCCACTACCCGGGGTCTCTCTCCAGCACCAGCAGCAGTGGCGCGCCCTCTTGGTGG 772
 QY 1182 AGGGGACCCGGTCTCCACACACACATCGCGCCCTAGAGAGCGCAGCATCTGGAGCA 1241
 Db 773 AGGGGACCCGGTCTCCACACACACATCGCGCCCTAGAGAGCGCAGCATCTGGAGCA 832
 QY 1242 AAGAGAAGGATAAACAAGAGGACACCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGG 1301
 Db 833 AAGAGAAGGATAAACAAGAGGACACCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGG 892
 QY 1302 CTCGGTAGGTGAAAAGCGAG 1321
 Db 893 CTCGGTAGGTGAAAAGCGAG 912

RESULT 14
 US-10-241-220-45
 ; Sequence 45, Application US/10241220
 ; Publication No. US20030148408A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Frantz, Gretchen
 ; APPLICANT: Hillan, Kenneth J.
 ; APPLICANT: Phillips, Heidi
 ; APPLICANT: Polakis, Paul
 ; APPLICANT: Spencer, Susan
 ; APPLICANT: Williams, P. Mickey
 ; APPLICANT: Wu, Thomas
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
 ; TREATMENT OF TUMOR
 ; FILE REFERENCE: P501081-US
 ; CURRENT APPLICATION NUMBER: US/10/241,220
 ; CURRENT FILING DATE: 2002-12-13
 ; NUMBER OF SEQ ID NOS: 120
 ; SEQ ID NO 45
 ; LENGTH: 806
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-10-241-220-45

Query Match 58.1%; Score 767.4; DB 15; Length 806;
 Best Local Similarity 58.0%; Pred. No. 2.8e-166;
 Matches 777; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
 QY 499 TTGTTCCAGCATGAGATCAGGAGCTGGAGTTGTTTCAGATCATCATCGTGGT 558
 Db 14 TCTCCTCGAAACAGGCAATGGCGAGCTGGAGTTGTTTCAGATCATCATCGTGGT 73
 QY 559 GGTGATGATGGTATGGTGGTGTGATCAGCTGCTGAGCCACTACAAGCTGTCTGC 618
 Db 74 GGTGATGATGGTATGGTGGTGTGATCAGCTGCTGAGCCACTACAAGCTGTCTGC 133
 QY 619 ACGGTCTTTCATAGCCGGACACAGCAGGGGGCGGAGAGAGATGCCCTGTCTCAGA 678
 Db 134 ACGGTCTTTCATAGCCGGACACAGCAGGGGGCGGAGAGAGATGCCCTGTCTCAGA 193
 QY 679 AGGATGCTGTGGCCCTCGAGAGACAGATGTTCAGGCAACGGATCCAGAGCCGAGGT 738
 Db 194 AGGATGCTGTGGCCCTCGAGAGACAGATGTTCAGGCAACGGATCCAGAGCCGAGGT 253
 QY 739 CTACGCCCGGCTCGGCCACCGCAGCGCTGGCGCGCCCTTCGCCCGAGCGGAGCG 798
 Db 254 CTACGCCCGGCTCGGCCACCGCAGCGCTGGCGCGCCCTTCGCCCGAGCGGAGCG 313
 QY 799 CTTCCACCGCTTCAGCCCACTATCGTACCTGACAGCAGATCGACCTGCCACCCAC 858
 Db 314 CTTCCACCGCTTCAGCCCACTATCGTACCTGACAGCAGATCGACCTGCCACCCAC 373
 QY 859 CATCTCGTGTGACAGCGGGAGGAGCCCGCCACCTACAGGGGCCCTGACCCCTCCAGCT 918

Db 374 CATCTCGTGTACAGCGGGAGGAGCCCGCCACCTACAGGGGCCCTTCACCTCCAGCT 433
 QY 919 TCGGGACCCCGAGCAGCAGTGGAACTGAACCGGAGTTCGTTGCGCGCAGCCCGCAACAG 978
 Db 434 TCGGGACCCCGAGCAGCAGTGGAACTGAACCGGAGTTCGTTGCGCGCAGCCCGCAACAG 493
 QY 979 AACCATCTTGACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1038
 Db 494 AACCATCTTGACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 553
 QY 1039 CAGTAACTCGGGCATCAGCGCCAGTGTCTAGCGGAGCGGGCGGCGCATGAGGGGGCGGCC 1098
 Db 554 CAGTAACTCGGGCATCAGCGCCAGTGTCTAGCGGAGCGGGCGGCGCATGAGGGGGCGGCC 613
 QY 1099 GCCCACTTACAGCAGGTATCGGCCACTACCGGGGTCTCTTTCAGCAGCAGCAGAG 1158
 Db 614 GCCCACTTACAGCAGGTATCGGCCACTACCGGGGTCTCTTTCAGCAGCAGCAGAG 673
 QY 1159 CAGTGGGCGGCCCTCTTGTCTGAGGGAGCGCCCGCTCCACACACACATCGCGCCCT 1218
 Db 674 CAGTGGGCGGCCCTCTTGTCTGAGGGAGCGCCCGCTCCACACACACATCGCGCCCT 733
 QY 1219 AGAGAGCGCAGCCATCTGGAGCAAGAGAGATAAAGAGAGAGAGAGAGAGAGAGAG 1278
 Db 734 AGAGAGCGCAGCCATCTGGAGCAAGAGAGATAAAGAGAGAGAGAGAGAGAGAGAG 793
 QY 1279 TCCCCAGGGGGC 1291
 Db 794 TCCCCAGGAAGGC 806

RESULT 15
 US-10-390-045-2
 ; Sequence 2, Application US/10390045
 ; Publication No. US20030170713A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SRIVASTAVA, SHIV
 ; APPLICANT: MOUL, JUDD W.
 ; APPLICANT: XU, LINDA L.
 ; APPLICANT: SEGAWA, TAKEHIKO
 ; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED
 ; FILE REFERENCE: P04995.0057-00000
 ; CURRENT APPLICATION NUMBER: US/10/390,045
 ; CURRENT FILING DATE: 2003-03-18
 ; PRIOR APPLICATION NUMBER: US/09/769,482
 ; PRIOR FILING DATE: 2001-01-26
 ; PRIOR APPLICATION NUMBER: 60/178,772
 ; PRIOR FILING DATE: 2000-01-28
 ; PRIOR APPLICATION NUMBER: 60/179,045
 ; PRIOR FILING DATE: 2000-01-31
 ; NUMBER OF SEQ ID NOS: 67
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 759
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-390-045-2

Query Match 57.2%; Score 755.6; DB 15; Length 759;
 Best Local Similarity 99.7%; Pred. No. 1.3e-163;
 Matches 757; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 518 ATCAGGAGCTGGAGTTTGTTCAGATCATCATCATCATCATCATCATCATCATCATCAT 577
 Db 1 ATGGCGGAGCTGGAGTTTGTTCAGATCATCATCATCATCATCATCATCATCATCATCAT 60
 QY 578 GTGGTGTACCTGCTGCTGAGCCACTACAGCTGTCTGACCGGTCTCTTATCAGCCGG 637
 Db 61 GTGGTGTACCTGCTGCTGAGCCACTACAGCTGTCTGACCGGTCTCTTATCAGCCGG 120
 QY 638 CACAGCCAGGGCGGAGGAGAGATGCTCTGTCTCCTCAGAGAGATGCTGTGGCCCTCG 697

```
Db 121 CACACCCAGGGCGGAGAGAGAGATGCCCTGTCTCTCAGAGGATGCTGTGCCCCCTCG 180
QY GAGAGCACAAGTGTACGGAACCGGAATCCAGAGCGCAGGTCTACGCCCCCGCTCGCCCC 757
Db 181 GAGAGCACAAGTGTACGGAACCGGAATCCAGAGCGCAGGTCTACGCCCCCGCTCGCCCC 240
QY 759 ACCGACCGCTGGCCGTCGCGCCCTTCGCCAGAGGGGAGCGCTTCCACCGCTTCCAGCCC 817
Db 241 ACCGACCGCTGGCCGTCGCGCCCTTCGCCAGAGGGGAGCGCTTCCACCGCTTCCAGCCC 300
QY 818 ACCTATCCGTACCTCCAGCAGAGATCGACCTGCGCACCCACCATCTCGCTCTCAGACGGG 877
Db 301 ACCTATCCGTACCTCCAGCAGAGATCGACCTGCGCACCCACCATCTCGCTCTCAGACGGG 360
QY 878 GAGGAGCCCCCACCCTACAGGGCCCTCGACCCCTCCAGCTTCGGGACCCCGAGCAGCAG 937
Db 361 GAGGAGCCCCCACCCTACAGGGCCCTCGACCCCTCCAGCTTCGGGACCCCGAGCAGCAG 420
QY 938 CTGGAACCTGAACCGGGAGTGGTGGCGCACCCCCCAACAGAACCATCTTCGACAGTGAC 997
Db 421 CTGGAACCTGAACCGGGAGTGGTGGCGCACCCCCCAACAGAACCATCTTCGACAGTGAC 480
QY 998 CTGATGGATAGTCCAGGCTGGGCGGCCCTGCCCCCGCCCGCCAGCAGTAAGTCCGGCATCAGC 1057
Db 481 CTGATGGATAGTCCAGGCTGGGCGGCCCTGCCCCCGCCCGCCAGCAGTAAGTCCGGCATCAGC 540
QY 1058 GCCACGTGTACCGCAGCGGGCGGCATGGAGGGGCGCGCCACCTACAGCGAGGTC 1117
Db 541 GCCACGTGTACCGCAGCGGGCGGCATGGAGGGGCGCGCCACCTACAGCGAGGTC 600
QY 1119 ATCGGCACTACCGGGGTCTCTCTTCCAGCACCAGCAGAGAGTGGGCGCCCTCTCTTG 1177
Db 601 ATCGGCACTACCGGGGTCTCTCTTCCAGCACCAGCAGAGAGTGGGCGCCCTCTCTTG 660
QY 1178 CTGGAGGGGACCCCGGTCCACACACACATCGCGGCCCTAGAGAGCGCAGCCATCTGG 1237
Db 661 CTGGAGGGGACCCCGGTCCACACACACATCGCGGCCCTAGAGAGCGCAGCCATCTGG 720
QY 1238 AGCAAGAGAGGATAAAGAGAAAGGACACCCCTCTCTAG 1276
Db 721 AGCAAGAGAGGATAAAGAGAAAGGACACCCCTCTCTAG 759
```

Search completed: May 12, 2004, 08:24:15
Job time : 669.311 secs

AL517150 AL517150
AL578575 AL578575
BQ641849 AGENCOURT
BM922276 AGENCOURT

SUMMARIES
LINE LOCAL

```

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DA008CA12NP1.
Location/Qualifiers
1. .1201
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
   /clone="CS0DA008YB23"
   /tissue_type="NEUROBLASTOMA"
   /clone_lib="Homo sapiens NEUROBLASTOMA"
   /note="Vector; pCMVSPORT 6; 1st strand cDNA was primed
with a Ndu1-oligo(3d) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

```

ORIGIN

Query Match		65.6%;	Score 866;	DB 9;	Length 1201;
Best Local Similarity		88.9%;	Pred. No. 7.8e-111;		
Matches 926;		Conservative 23;	Mismatches 91;	Indels 2;	Gaps 2;
Qy	280	GCCCGGAGACCGGGGCGCTCCCGCGCGCGCTCTCTGATCGGGGCCCCAGCT	339	1157	SCCCSGSCCCCGTTTKSGGCCCCSGSCCGGGSCGGGGCTTCCCCCGGGCTCBGG
Db	1157	SCCCSGSCCCCGTTTKSGGCCCCSGSCCGGGSCGGGGCTTCCCCCGGGCTCBGG	1098	340	CCGGGCGCGCGCGAGACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
Qy	340	CCGGGCGCGCGCGAGACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	399	1097	TKGGGCGCCGCTCCGGGSGRGGASCCCGCGGAAAAAATAAAWCSGRMAAGGC
Db	1097	TKGGGCGCCGCTCCGGGSGRGGASCCCGCGGAAAAAATAAAWCSGRMAAGGC	1038	400	CGCGCGCGCTGCATGCACCGCTGTATGGGGGTCAACAGCACCGCGCGCGCGCGCG
Qy	400	CGCGCGCGCTGCATGCACCGCTGTATGGGGGTCAACAGCACCGCGCGCGCGCGCG	459	1037	CGCCSGSCMGTCWATCACCGCTGTGTGGGTCAACAGCACCGCGCGCGCGCGCG
Db	1037	CGCCSGSCMGTCWATCACCGCTGTGTGGGTCAACAGCACCGCGCGCGCGCGCG	979	460	GCAGCCCAATGTCCTGCACGTGCAACTGCACAAACGCTCTTTGTTCCAGAGCATGGAGAT
Qy	460	GCAGCCCAATGTCCTGCACGTGCAACTGCACAAACGCTCTTTGTTCCAGAGCATGGAGAT	519	978	GCAGCCCAATTTCTCTGCACGTGCCATCTGCAACGCTTTTGTTCAGAGCATGGAGAT
Db	978	GCAGCCCAATTTCTCTGCACGTGCCATCTGCAACGCTTTTGTTCAGAGCATGGAGAT	919	520	CACGGAGCTGGAGTTTGTTCAGATCATCATCTGTGTGTGTATGATGATGTGTGT
Qy	520	CACGGAGCTGGAGTTTGTTCAGATCATCATCTGTGTGTGTATGATGATGTGTGT	579	918	CACGGAGCTGGAGTTTGTTCAGATCATCATCTGTGTGTGTATGATGTGTGTGT
Db	918	CACGGAGCTGGAGTTTGTTCAGATCATCATCTGTGTGTGTATGATGTGTGTGT	859	580	GGTGATCAGTGCCTGCTGAGCCATACAAGCTGCTGCACGGTCCCTTCATCAGCCGCA
Qy	580	GGTGATCAGTGCCTGCTGAGCCATACAAGCTGCTGCACGGTCCCTTCATCAGCCGCA	639	858	GGTGATCAGTGCCTGCTGAGCCATACAAGCTGCTGCACGGTCCCTTCATCAGCCGCA
Db	858	GGTGATCAGTGCCTGCTGAGCCATACAAGCTGCTGCACGGTCCCTTCATCAGCCGCA	799	640	CAGCCAGGGGCGGAGGAGAGATGCCCTGTCTTCAGAGGATGCTGTGGCCCTCGGA
Qy	640	CAGCCAGGGGCGGAGGAGAGATGCCCTGTCTTCAGAGGATGCTGTGGCCCTCGGA	699	798	CAGCCAGGGGCGGAGGAGAGATGCCCTGTCTTCAGAGGATGCTGTGGCCCTCGGA
Db	798	CAGCCAGGGGCGGAGGAGAGATGCCCTGTCTTCAGAGGATGCTGTGGCCCTCGGA	739	700	GAGCACAGTGTACGGCAACGGAAATCCGACAGCGCAGTGTACGGCCCGCTTCGGCCAC
Qy	700	GAGCACAGTGTACGGCAACGGAAATCCGACAGCGCAGTGTACGGCCCGCTTCGGCCAC	759	738	GAGCACAGTGTACGGCAACGGAAATCCGACAGCGCA-GTCTACGGCCCGCTTCGGCCAC
Db	738	GAGCACAGTGTACGGCAACGGAAATCCGACAGCGCA-GTCTACGGCCCGCTTCGGCCAC	680	760	CGACCGCTGCGCGTGGCGCCCTTCGCCACGGGAGCGCTTCACACGCTTCACAGCCAC
Qy	760	CGACCGCTGCGCGTGGCGCCCTTCGCCACGGGAGCGCTTCACACGCTTCACAGCCAC	819	679	CGACCGCTGCGCGTGGCGCCCTTCGCCACGGGAGCGCTTCACACGCTTCACAGCCAC
Db	679	CGACCGCTGCGCGTGGCGCCCTTCGCCACGGGAGCGCTTCACACGCTTCACAGCCAC	620	820	CTATCCGTACCTGCAGCACGAGATCGACCTCGCACCCACCATCTCTGCTGCAGACGGGA
Qy	820	CTATCCGTACCTGCAGCACGAGATCGACCTCGCACCCACCATCTCTGCTGCAGACGGGA	879	619	CTATCCGTACCTGCAGCACGAGATCGACCTCGCACCCACCTTCTCGCTGTACAGACGGGA
Db	619	CTATCCGTACCTGCAGCACGAGATCGACCTCGCACCCACCTTCTCGCTGTACAGACGGGA	560	880	GGAGCCCCCAACCTTACGAGGCCCCCTGCACCCCTCAGCTTCGGGACCCCGAGACGACCT
Qy	880	GGAGCCCCCAACCTTACGAGGCCCCCTGCACCCCTCAGCTTCGGGACCCCGAGACGACCT	939	559	GGAGCCCCCAACCTTACGAGGCCCCCTGCACCCCTCAGCTTCGGGACCCCGAGACGACCT
Db	559	GGAGCCCCCAACCTTACGAGGCCCCCTGCACCCCTCAGCTTCGGGACCCCGAGACGACCT	500	940	GGAACTGAACCGGGAGTCGGTGGCGGCACCCCCAAACAGAACCACTTCTTCGACAGTGACCT
Qy	940	GGAACTGAACCGGGAGTCGGTGGCGGCACCCCCAAACAGAACCACTTCTTCGACAGTGACCT	999	499	GGGACTGAACCGGGAGTCGGTGGCGGCACCCCCAAACAGAACCACTTCTTCGACAGTGACCT
Db	499	GGGACTGAACCGGGAGTCGGTGGCGGCACCCCCAAACAGAACCACTTCTTCGACAGTGACCT	440	1000	GATGANTAGTGCAGGCTGGGCGGCCCTGTGCCCCCCCCAGCAGTAACCTCGGCATCAGCGC
Qy	1000	GATGANTAGTGCAGGCTGGGCGGCCCTGTGCCCCCCCCAGCAGTAACCTCGGCATCAGCGC	1059	439	GATGANTAGTGCAGGCTGGGCGGCCCTGTGCCCCCCCCAGCAGTAACCTCGGCATCAGCGC
Db	439	GATGANTAGTGCAGGCTGGGCGGCCCTGTGCCCCCCCCAGCAGTAACCTCGGCATCAGCGC	380		

[illegible]

QY 1064 TGCTACGGCAGCGGGCGGCATGAGGGCGCGCCGCCACCTACAGGAGGTCTATCGC 1123
 Db 541 TGCTACGGCAGCGGGCGGCATGAGGGCGCGCCGCCACCTACAGGAGGTCTATCGC 600
 QY 1124 CACTACCCGGGGTCTCTCTCCAGCACCAGCAGCAGTGGCGCCCTCTCTTGTGGAG 1183
 Db 601 CACTACCCGGGGTCTCTCTCCAGCACCAGCAGCAGTGGCGCCCTCTCTTGTGGAG 660
 QY 1184 GGGACCCGGCTCCACACACACATCGCGCCCTAGAGCGCAGCCATCTGGAGCAA 1243
 Db 661 GGGACCCGGCTCCACACACACATCGCGCCCTAGAGCGCAGCCATCTGGAGCAA 720
 QY 1244 GAGAAGATAAAGAGAGACACCTCTCTAGGGTCCCGGGGGCGCGGCTCGGG 1301
 Db 721 GAGAAGATAAAGAGAGACACCTCTCTAGGGTCCCGGGGGCGCGGCTCGGG 778

RESULT 4

BM922276
 LOCUS
 DEFINITION BM922276 1046 bp mRNA linear EST 12-MAR-2002
 5', mRNA sequence.

ACCESSION BM922276.1 GI:19372655

VERSION

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

NIH-MGC <http://mgs.nhl.nih.gov/>;
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clones distributed: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM12791 row: n column: 06
 High quality sequence stop: 671.

FEATURES

source

1..1046

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5754437"

/lab_host="DH10B"

/clone_lib="NIH_MGC_115"

/note="Organ: pooled brain, lung, testis; Vector:
 pCMV-SF0K16; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
 source anonymous pool of 6 male brains, age range 23-27; 1
 male lung, age 27; and 1 male testis, age 69. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.8 kb,
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research genetics tracking code
 021. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 58.0%; Score 766; DB 12; Length 1046;
 Best Local Similarity 95.8%; Pred. No. 61e-97;
 Matches 819; Conservative 0; Mismatches 32; Indels 4; Gaps 3;
 QY 347 CCGGCGGAGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 406
 Db 1 CCGGCGGAGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60

QY 407 CCGTCCATGACCGCTTGATGGGGTCAACAGCACCGCGCGCGCGCGCGCGCGCGCG 466
 Db 61 CCGTCCATGACCGCTTGATGGGGTCAACAGCACCGCGCGCGCGCGCGCGCGCGCG 120
 QY 467 AATGTCTCTGACGTGCAACTGCAAAACGCTCTTTGTTCCAGAGCATGGAGATCA 526
 Db 121 AATGTCTCTGACGTGCAACTGCAAAACGCTCTTTGTTCCAGAGCATGGAGATCA 180
 QY 527 CTGGAGTTTCTCAGATCATCATCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 586
 Db 181 CTGGAGTTTCTCAGATCATCATCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 240
 QY 587 ACGTGCCTGTGAGCCACATCAAGCTGTCTGCAAGCTCTTTCATCAGCCGCA 646
 Db 241 ACGTGCCTGTGAGCCACATCAAGCTGTCTGCAAGCTCTTTCATCAGCCGCA 300
 QY 647 GGGCGGAGGAGAGAGATGCGCTCTCAGAGGATGCTGTGGCCCTCGGAGGACACA 706
 Db 301 GGGCGGAGGAGAGAGATGCGCTCTCAGAGGATGCTGTGGCCCTCGGAGGACACA 360
 QY 707 GTGTTCAGGCAACGGAATCCAGAGCGCGAGTCTACGCGCCCGCTCGGCCAC 766
 Db 361 GTGTTCAGGCAACGGAATCCAGAGCGCGAGTCTACGCGCCCGCTCGGCCAC 420
 QY 767 CTGGCGTGTGCGCCCTTGGCCGAGCGGAGCGCTTCCACGGCTTCCAGCCAC 826
 Db 421 CTGGCGTGTGCGCCCTTGGCCGAGCGGAGCGCTTCCACGGCTTCCAGCCAC 480
 QY 827 TACTCTGAGCAGCAGATCGACCTCGCCACCCACCATCTCGCTGTGACAGCGG 886
 Db 481 TACTCTGAGCAGCAGATCGACCTCGCCACCCACCATCTCGCTGTGACAGCGG 540
 QY 887 CCACCTTACAGGGCGCCCTGACCTTCAGCTTCGGGACCCCGAGCAGCAGTGG 946
 Db 541 CCACCTTACAGGGCGCCCTGACCTTCAGCTTCGGGACCCCGAGCAGCAGTGG 600
 QY 947 AACCGGAGTCGGTGGCGCACCCCAACAGAACCATCTTCGACAGTACCTGAT 1006
 Db 601 AACCGGAGTCGGTGGCGCACCCCAACAGAACCATCTTCGACAGTACCTGAT 660
 QY 1007 AGTCCAGGTCGGGCGCCCTGCGCCCGCAGCAGTAACCTCGGGCATCAGGCCA 1066
 Db 661 AGTCCAGGTCGGGCGCCCTGCGCCCGCAGCAGTAACCTCGGGCATCAGGCCA 720
 QY 1067 TACGGCAGCGCGCG-GCGCATGGAGGGCGCGCG-CACCTACAGCGAGTCA 1124
 Db 721 TACGGCAGCG 780
 QY 1125 ACTA--CCCGGGTCTCTCTTCCAGCACCGAGAGCAGTGGCGCGCCCTCTT 1182
 Db 781 ACTAACCGGGGTCTCTCTTCCAGCACCGAGAGCAGTGGCGCGCGCGCGCGCG 840
 QY 1183 GGGACCCCGGCTCC 1196
 Db 841 CTGGAGGGGAACC 854

RESULT 5

AL543170

LOCUS

DEFINITION

AL543170

ACCESSION

AL543170

VERSION

AL543170.2

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 1201)

AUTHORS

Li, W.B., Gruber, C., Jesse, J., and Polayes, D.

TITLE

Full-length cDNA libraries and normalization

JOURNAL

Unpublished (2001)

AL543170 1201 bp mRNA linear EST 31-MAY-2003
 AL543170 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
 clone CSOD1002Y103 5-PRIME, mRNA sequence.

QY 1117 CATCGGCACACTACCGGGGTCCTCTTCCAGCACACGAGAGAGTGGGCGGCTCTCTT 1176
 Db 314 CATCGGCACACTACCGGGGTCCTCTTCCAGCACACGAGAGAGTGGGCGGCTCTCTT 255
 QY 1177 GCTGAGGGGACCGGGTCCACACACACATCGCCCTTCCAGAGCGAGCATCTG 1236
 Db 254 GCTGAGGGGACCGGGTCCACACACACATCGCCCTTCCAGAGCGAGCATCTG 195
 QY 1237 GAGCAAGAGAGGATTAACAGAAAGACACCTCTCTAGGGTCCCCAGGGGGCGGGC 1296
 Db 194 GAGCAAGAGAGGATTAACAGAAAGACACCTCTCTAGGGTCCCCAGGGGGCGGGC 135
 QY 1297 TGGGGCTGCTAGGTGAAAGGCGAG 1321
 Db 134 TGGGGCTGCTAGGTGAAAGGCGAG 110

RESULT 8
 AL558882 1007 bp mRNA linear EST 31-MAY-2003
 LOCUS
 DEFINITION Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
 Homo sapiens cDNA clone CS0DJ015F12 5-PRIME, mRNA sequence.
 ACCESSION AL558882
 VERSION AL558882.2 GI:31283015
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1007)
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On Feb 15, 2001 this sequence version replaced gi:12903838.
 COMMENT
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 9945.r For more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DJ015DC06QPI&cluster=9945.r. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DJ015DC06QPI.
 Location/Qualifiers
 1..1007
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DJ015F12"
 /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
 /cell_lines="JURKAT"
 /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
 10-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
 source

ORIGIN
 Query Match 54.2%; Score 715.4; DB 9; Length 1007;
 Best Local Similarity 99.1%; Pred. No. 6.6e-90;
 Matches 738; Conservative 2; Mismatches 3; Indels 2; Gaps 2;
 QY 577 GTGTGTATCATCGTCTGCTGAGCCACTACAAAGTGTGTGACGGTCTTTTCATGCGG 636
 Db 62 GGATGTGATCATCGTCTGCTGAGCCACTACAAAGTGTGTGACGGTCTTTTCATGCGG 121
 QY 637 GCACAGCCAGGGCGGAGAGAGATGCCCTGTCTCAGAGAGATCGCTGTGCGCCCTC 696
 Db 122 GCACAGCCAGGGCGGAGAGAGATGCCCTGTCTCAGAGAGATCGCTGTGCGCCCTC 181

QY 697 GGAGAGCACAGTGTTCAGGCAACGGAATCCAGAGCGCAGGCTCTACGCCCGCCTCGGCC 756
 Db 182 GGAGAGCACAGTGTTCAGGCAACGGAATCCAGAGCGCAGGCTCTACGCCCGCCTCGGCC 240
 QY 757 CACCGACCGCTTGGCGGCTTGGCGGCTTGGCGGCTTGGCGGCTTGGCGGCTTGGCGGCT 816
 Db 241 CACCGACCGCTTGGCGGCTTGGCGGCTTGGCGGCTTGGCGGCTTGGCGGCTTGGCGGCT 300
 QY 817 CACCTATCCGTACCTTGCAGCAGAGATCGAATCGAATCGAATCGAATCGAATCGAATCG 876
 Db 301 CACCTATCCGTACCTTGCAGCAGAGATCGAATCGAATCGAATCGAATCGAATCGAATCG 360
 QY 877 GGAGAGCGCCACACCTTACCAGGGCCCTTGCACCTTCCAGCTTCCGAGACCCGAGCAGCA 936
 Db 361 GGAGAGCGCCACACCTTACCAGGGCCCTTGCACCTTCCAGCTTCCGAGACCCGAGCAGCA 420
 QY 937 GCTGGAACCTGAACCGGGAGTGGTGGCGGACACCCCAACAGAAACATCTTTCGACAGTGA 996
 Db 421 GCTGGAACCTGAACCGGGAGTGGTGGCGGACACCCCAACAGAAACATCTTTCGACAGTGA 480
 QY 997 CTTGATGATAGTGCAGGCTGGGGGGCCCTTGCACGACACGACGAGAGTGGGCGGCTCT 1056
 Db 481 CTTGATGATAGTGCAGGCTGGGGGGCCCTTGCACGACACGACGAGAGTGGGCGGCTCT 540
 QY 1057 CGCCACGCTGCTACGGCAGCGGGCGGCGCATGGAGGGGGCGGCGGCGGCGGCGGCGGCT 1116
 Db 541 CGCCACGCTGCTACGGCAGCGGGCGGCGCATGGAGGGGGCGGCGGCGGCGGCGGCGGCT 600
 QY 1117 CATCGGCCACTACCGGGGTCCTCTTCCAGCACACGAGAGAGTGGGCGGCTCTCTT 1176
 Db 601 CATCGGCCACTACCGGGGTCCTCTTCCAGCACACGAGAGAGTGGGCGGCTCTCTT 660
 QY 1177 GCTGAGGGGAGCCCGGCTCCACACACACATCGCGGCGGCGGCGGCGGCGGCGGCTCTG 1236
 Db 661 GCTGAGGGGAGCCCGGCTCCACACACACATCGCGGCGGCGGCGGCGGCGGCGGCTCTG 719
 QY 1237 GAGCAAGAGAGAGTAAACAGAAAGACACCTCTCTAGGGTCCCCAGGGGGCGGGC 1296
 Db 720 GAGCAAGAGAGAGTAAACAGAAAGACACCTCTCTAGGGTCCCCAGGGGGCGGGC 779
 QY 1297 TGGGGCTGCTAGGTGAAAGGCGAG 1321
 Db 780 TGGGGCTGCTAGGTGAAAGGCGAG 804

RESULT 9
 BX362396/c 874 bp mRNA linear EST 05-MAY-2003
 LOCUS
 DEFINITION BX362396 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
 Homo sapiens cDNA clone CS0DJ014YN15 3-PRIME, mRNA sequence.
 ACCESSION BX362396
 VERSION BX362396.1 GI:30378625
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 874)
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 9945.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DJ014CG08NP1&cluster=9945.r. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0DJ014CG08NP1.

FEATURES

```

1. 874
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D0104XN15"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN

Query Match	54.1%	Score 715	DB 13	Length 874
Best Local Similarity	96.2%	Pred. No. 7.3e-90		
Matches 731	Conservative 16	Mismatches 11	Indels 2	Gaps 2
QY	563	ATGATGGTGATGGTGGTGGTATC-ACGTCCCTGCTGAGCCACTCAAGCTGTCTGACG	621	
DB	874	ATGATGGTGATGGTGGTGGTATC-ACGTCCCTGCTGAGCCACTCAAGCTGTCTGACG	815	
QY	622	GTCTTTCATCAGCCGGCAGCAGCCAGGGGCGGAGAGAGAGATGCCCTGTCTCAGAAGG	681	
DB	814	GTCTTTCATCAGCCGGCAGCAGCCAGGGGCGGAGAGAGATGCCCTKTCVTCAGAAGG	755	
QY	682	ATGCCCTGTGGCCCTCGGAGGACACAGTGTGAGCAACGGAAATCCAGAGCCGCACTTA	741	
DB	754	ATGCTYTKTGGCCCTCGGAGGACACAGTGTGAGCAACGGAAATCCAGAGCCGCA-ATYTA	696	
QY	742	CGCCCCGGCTCGGGCCACCCAGCCGCTGGCCGCTTGGCCGACGCGGAGCGCTT	801	
DB	695	CGCCCCGGCTTGGGCCACCCAGCCGCTTGGCCGCTTGGCCGACGCGGAGCGCTT	636	
QY	802	CCACCGCTTCAGGCCACCTATCCGTACTCTGACGACGAGATCGACTTGCCACCCACCAT	861	
DB	635	CCACCGCTTCAGGCCACCTATCCGTACTCTGACGACGAGATNGACNTGCCGCCACCAT	576	
QY	862	CTCGCTGTACAGCGGGAGGACCCACCCCTACGAGGCGCCCTGACACCTCCAGCTTCG	921	
DB	575	CTCGCTGTACAGCGGGAGGACCCACCCCTACGAGGCGCCCTGACACCTCMAGTTTCG	516	
QY	922	GGACCCCGAGCAGCAGCTGGAACTGAACCGGGAGTCGGTGGCGGACCCGCCAAAACAGAAC	981	
DB	515	GGACCCCGAGWAGCAGCTGGAACCTGAACCGGGAGTCGGTGGCGGACCCGCCAAAACAGAAC	456	
QY	982	CATCTTTGACAGTGACTGATGATAGTGTGCGAGCTGGCGGGCCCTGCGCCGCCAGCAG	1041	
DB	455	CATCTTTGACAGTGACTGATGATGATGCGAGCTGGCGGGCCCTGCGCCGCCAGCAG	396	
QY	1042	TAACTCGGGCATCAGCGCCAGCTGCTACGGCAGCGCGGGCGCATGGAGGGGCCGCCGCC	1101	
DB	395	TAACTCGGGCATCAGCGCCAGCTGCTACGGCAGCGCGGGCGCATGGAGGGGCCGCCGCC	336	
QY	1102	CACCTACAGCGAGGTTCATCGGCCACTACCGGGGTCCTCTTCAGCACCAGCAGCAG	1161	
DB	335	CACCTACAGCGAGGTTCATCGGCCACTACCGGGGTCCTCTTCAGCACCAGCAGCAG	276	
QY	1162	TGGGCGCCCTTCCTTGTCTGGAGGGGACCCGGCTCCACACACACATCGCGGCCCTAGA	1221	
DB	275	TGGGCGCCCTTCCTTGTCTGGAGGGGACCCGGCTCCACCAACACATCGCGGCCCTAGA	216	
QY	1222	GAGCGCAGCCATCTGGAGCAAGAGAGGATTAACAGAAAGGACACCTCTCTTAGGGTCC	1281	
DB	215	GAGCGCAGCCATCTGGAGCAAGAGAGGATTAACAGAAAGGACACCTCTCTTAGGGTCC	156	
QY	1282	CCAGGGGGCGCGGGCTGGGGCTGGCTAGGTGAAAGGAC	1321	
DB	155	CCAGGGGGCGCGGGCTGGGGCTGGCTAGGTGAAAGGAC	116	

RESULT 10
BU539219

```

BUS539219      945 bp      mRNA      linear      EST 13-SEP-2002
AGENCOURT 10215265 NIH_MGC.107 Homo sapiens cDNA clone
IMAGE:6569922 5', mRNA sequence.

BUS539219
BUS539219.1   GI:22849660
EST.

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 945)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: AICC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCN2757 row: p column: 18
High quality sequence stop: 663.

```

FEATURES	SOURCE
1. High Accuracy: The model achieves a high accuracy rate, indicating its effectiveness in predicting the target variable.	Model Performance Metrics
2. Robustness: The model demonstrates robustness against various input variations and noise, ensuring reliable predictions.	Model Performance Metrics
3. Interpretability: The model's predictions are interpretable, allowing users to understand the underlying factors influencing the outcome.	Model Performance Metrics
4. Scalability: The model is scalable, capable of handling large volumes of data and complex tasks efficiently.	Model Performance Metrics
5. Real-time Processing: The model is designed for real-time processing, enabling immediate decision-making based on incoming data.	Model Performance Metrics
6. Customizability: The model is highly customizable, allowing users to tailor it to specific requirements and datasets.	Model Performance Metrics
7. Integration: The model integrates seamlessly with existing systems and workflows, facilitating easy adoption and implementation.	Model Performance Metrics
8. Security: The model adheres to strict security protocols, ensuring that data is protected and handled responsibly.	Model Performance Metrics
9. Compliance: The model complies with relevant regulatory requirements, ensuring legal and ethical standards are met.	Model Performance Metrics
10. Support: Comprehensive support and documentation are provided to assist users in maximizing the model's performance.	Model Performance Metrics

```

1. 945
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:656922"
/tissue_type="adenocarcinoma, cell line"
/lab_host="PH108 (phage-resistant)"
/clone_lib="NIH MGC 107"
/notes="Organ: breast; Vector: pDPB7; Site 1: EcoRI;
Site 2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites, using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH MGC Library."

```

ORIGIN

Query Match	52.3%;	Score	691.2;	DB 13;	Length	945;			
Best Local Similarity	93.7%;	Prod. No.	1.5e-86;						
Matches	764;	Conservative	0;	Mismatches	44;	Indels	7;	Gaps	4;
Qy	499	TTTGTTCAGAGCATGGAGATCAGCGAGCTGGAGTTTGTTCAGATCATCATCATCGTGGT	558						
Db	65	TCTCTGGAACACAGGCAATGGCGAGCTGGAGTTTGTTCAGATCATCATCATCGTGGT	124						
Qy	559	GCTGATGATGFGATGFGTGFATCACGTGCCTGCTGAGCCACTCAAGCTGTCTGC	618						
Db	125	CGTGATGATGGTATGGTGGTGATCACTGCTGCTGAGCCACTCAAGCTGTCTGC	184						
Qy	619	AGGTCCTTCATCAGCCGGCACAGCAGAGGGCGGAGAGAAAGATGCCCTCTCTCAGA	678						
Db	185	ACGGTCCTTCATCAGCTGGCACAGCAGGGGGCGAGAGAGATGCCCTCTCTCAGA	244						
Qy	679	AGGATGCTGTGGCCCTCGGAGAGACAGTGTCAAGCAACGGAAATCCCAAGACCCGAGGT	738						
Db	245	AGGATGCTGTGGCCCTCGGAGAGACAGTGTCAAGCAACGGAAATCCCAAGACCCGAGGT	304						
Qy	739	CTACGCCCGCCTGGGCCCAACGACCGCCTGGCCGTGGCCCTTCGCCACGGGGAGCG	798						
Db	305	CTACGCCCGCCTGGGCCCAACGACCGCCTGGCCGTGGCCCTTCGCCACGGGGAGCG	364						
Qy	799	CTTCACACGCCTTCCAGCCCACTATCGGTACTCTGCACGACGAGATCGACTCTCCACCCAC	858						

Db 365 CTTCCACCGCTTCCAGCCACCTATCCGTACCTGTCAGCAGAGATCGACCTGCGGCCAC 424
 Qy 859 CATCTCGTGTGTCAGAGCGGGAGAGCCCTACCTACAGGCGCCCTGCACCTCCAGCT 918
 Db 425 CATCTCGTGTGTCAGAGCGGGAGAGCCCTACCTACAGGCGCCCTGCACCTCCAGCT 484
 Qy 919 TCGGACCCCGAGCAGCAGCTGGAACCTGAAACCGGGAGTCGTGCGCGCACCCCAACAG 978
 Db 485 TCGGACCCCGAGCAGCAGCTGGAACCTGAAACCGGGAGTCGTGCGCGCACCCCAACAG 544
 Qy 979 AACCATCTTCAGCAGTGCCTGATGATAGTCGACAGGTGCGGCGCCCTGCGCCCGCAG 1038
 Db 545 AACCATCTTCAGCAGTGCCTGATGATAGTCGACAGGTGCGGCGCCCTGCGCCCGCAG 604
 Qy 1039 CAGTAACCTCGGCGCATCAGCGCCACCTGCTATCGGCAGCGCGCGGCATGAGAGGCGCGCC 1098
 Db 605 CAGTAACCTCGGCGCATCAGCGCCACCTGCTATCGGCAGCGCGCGGCATGAGAGGCGCGCC 664
 Qy 1099 GCCCACCCTACAGCAGGTGATCGGCCACTACCCGGGTCTCTCTTCAGCAGCAGCAGAG 1158
 Db 665 GCCCACCCTACAGCAGGTGATCGGCCACTACCCGGGTCTCTCTTCAGCAGCAGCAGAG 724
 Qy 1159 CAGTGGGCGCGCCT-CCTTGTCTGGAGGGAGCCCGGCTCCACACACACATCGCG-CCC 1216
 Db 725 CAGTGGGCGCGCCTCTCTTGTCTGGAGGGAGCCCGGCTCCACACACACATCGTGCCCC 784
 Qy 1217 CTAGAGAGCGCAG-CATCTGAGCAAGAGAGGATAA-----CAGAAAGGACACCCCTC 1271
 Db 785 CTAAGAGCGCAGCCCATCTGAGCAAGAGAGGATAAACCCGAAAGGACACCCCTTC 844
 Qy 1272 TCTAGGTCCTCCAGAGGGCGCGGCTGGGGCTGGCG 1306
 Db 845 TCTAGTGTCTCCCGAGGGGGGCGCGGCTGGG 879

RESULT 11

BU602918
 LOCUS AGENCORET.10016502 NIH_MGC_142 Homo sapiens cDNA clone
 DEFINITION IMAGE:6497853 5', mRNA sequence.

BU602918
 VERSION BU602918.1 GI:23254677

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 850)

NIH-MGC <http://mgs.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: NCI

cDNA Library Preparation: Michael Brownstein Laboratory

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Place: LNC2679 row: i column: 22

High quality sequence stop: 499.

FEATURES

source

1..850

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6497853"

/tissue_type="mixed (pool of 40 RNAs)"

/lab_host="DH10B (Tl-phage-resistant)"

/clone_lib="NIH MGC 142"

/notes="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc);

Site 2: SfiI (ggcgctcgcc); Double-stranded cDNA was

prepared from a pool of 40 cell line polyA+ RNAs (bladder - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon - 4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%, kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were used in cloning as follows:

5'-AAGCATGTTATCAACGACGAGTGGCCATTACGGCGGG-3' and
 5'-ATTCTAGAGCGAGCGCGGCACATG-dt(30)NN-3'. Full-length
 enriched library was constructed using the Clontech
 Creator SMART kit and size-selected to contain the >0.5 kb
 size fraction (other fractions present in NIH MGC 141).
 Library created in the laboratory of M. Brownstein (NIH,
 NIH). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 52.0%; Score 686.6; DB 13; Length 850;
 Query Local Similarity 96.9%; Pred. No. 6.5e-86;
 Matches 743; Conservative 0; Mismatches 19; Indels 5; Gaps 4;

Qy 559 GGTGATGATGTTGATGTTGTTGATCAGTGTCTGCTGAGCCACTACAAGTGTCTGC 618
 Db 1 GGTGATGATGTTGATGTTGTTGATCAGTGTCTGCTGAGCCACTACAAGTGTCTGC 60
 Qy 619 ACCTGCTTTCATCAGCCGCGCAGCGGGGAGAGAGAGATGCTCTCTCAGA 678
 Db 61 ACCTGCTTTCATCAGCCGCGCAGCGGGGAGAGAGAGATGCTCTCTCAGA 120
 Qy 679 AGGATGCTGTGTCCTTCGAGAGACAGTGTTCAGGCAACGGAATCCGAGCGCGAGGT 738
 Db 121 AGGATGCTGTGTCCTTCGAGAGACAGTGTTCAGGCAACGGAATCCGAGCGCGAGGT 180
 Qy 739 CTAGCGCCCGCTCGGCGCCACCGCGCTGGCGGTGGCCCTTCGCCAGCGGGAGCG 798
 Db 181 CTAGCGCCCGCTCGGCGCCACCGCGCTGGCGGTGGCCCTTCGCCAGCGGGAGCG 240
 Qy 799 CTTCCACCGCTTCCAGCCCACTATCCGTAACCTGTCAGCAGAGATCGACTGCCACCCAC 858
 Db 241 CTTCCACCGCTTCCAGCCCACTATCCGTAACCTGTCAGCAGAGATCGACTGCCACCCAC 300
 Qy 859 CATCTCGTGTTCAGCGGGAGAGCGCCCACTACCGTTCAGGGCGCCCTGCACCTCCAGCT 918
 Db 301 CATCTCGTGTTCAGCGGGAGAGCGCCCACTACCGTTCAGGGCGCCCTGCACCTCCAGCT 360
 Qy 919 TCGGACCCCGAGCAGCAGCTGGAACTGAACCGGGAGTCGGTTCGGCGCACCCCAACAG 978
 Db 361 TCGGACCCCGAGCAGCAGCTGGAACTGAACCGGGAGTCGGTTCGGCGCACCCCAACAG 420
 Qy 979 AACCATCTTCAGCAGTGCACCTGATGATGTCAGCTGGGCGCGCCCTGCGCCCGCAG 1038
 Db 421 AACCATCTTCAGCAGTGCACCTGATGATGTCAGCTGGGCGCGCCCTGCGCCCGCAG 480
 Qy 1039 CAGTAACCTCGGCGCATCAGCGCCACCTGCTACCGCGAGCGCGCGGCATGAGAGGCGCGCC 1098
 Db 481 CAGTAACCTCGGCGCATCAGCGCCACCTGCTACCGCGAGCGCGGCATGAGAGGCGCGCC 540
 Qy 1099 GCCCACCCTACAGCAGGTGATTCGGCACAATAACCGGGGTCTCTTCAGCAGCAGCAGAG 1158
 Db 541 GCCCACCCTACAGCAGGTGATTCGGCACAATAACCGGGGTCTCTTCAGCAGCAGCAGAG 600
 Qy 1159 CAGTGGGCGCG-CCCTCCTCTCTG-AGGGGACCCCGCTCCACCACACACACATCGCGCCC 1216
 Db 601 CAGTGGGCGCGCCCTCTCTTCTGGAAGGGAGCCCGGTTCCACACACACACAAATCCGCCC 660
 Qy 1217 CTAGAGAGCGCAGCCA-TCTGGAGCAAGAGAGGATAAACAAGAGGACACCCCTCTCTA 1275
 Db 661 CTAAGAGCGCAGCCATTCTGGAGCAAGCAGAGATGATAAACCGAAGAGGAGCCCTCTCTA 720
 Qy 1276 GGGTCCCCCAGGGGGCC-CCGGCTGGGGCTGCGTAGTGAAAAGGCA 1320
 Db 721 GGGTCCCCCAGGGGGCCCGCGGCTGGGGGCTGTCTAGTGAAAAGGA 767


```

RESULT 12
AY419334
LOCUS
DEFINITION Homo sapiens TM6PAI gene, GSS 17-DEC-2003
GENOMIC SURVEY SEQUENCE.
ACCESSION AY419334
VERSION AY419334.1 GI:39775291
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 780)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Perriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
These sequences were made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..780
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>780
/gene="TM6PAI"
/locus_tag="HCM6856"
gene
ORIGIN
Query Match 49.3%; Score 651; DB 29; Length 780;
Best Local Similarity 86.2%; Pred. No. 5.7e-81;
Matches 551; Conservative 0; Mismatches 104; Indels 0; Gaps 0;
QY 522 CGGAGCTGGATTGTTTCAGATCATCATCATCGTGGTGGTGGTGGTGGTGGTGG 581
DB 26 CGGAGCTGGATTGTTTCAGATCATCATCATCGTGGTGGTGGTGGTGGTGG 85
QY 582 TGATCAGTGGCTGCTGAGCCACTACAGCTGTCTGCACGGTCTTTCATCAGCGGCACA 641
DB 86 TGATCAGTGGCTGCTGAGCCACTACAGCTGTCTGCACGGTCTTTCATCAGCGGCACA 145
QY 642 GCCAGGGCGGAGGAGAGAGATGCCCTGCTCCTCAGAGGATGCTGTGCCCTCGGAGA 701
DB 146 GCCAGGGCGGAGGAGAGAGATGCCCTGCTCCTCAGAGGATGCTGTGCCCTCGGAGA 205
QY 702 GCACAGTGTACGGAACCGGATCCAGCGCGCAGGCTACGCCGCCCTCGGCCACCG 761
DB 206 GCACAGTGTACGGAACCGGATCCAGCGCGCAGGATCCAGGNNNNNNNNNNNNNNNNNN 265
QY 762 ACCGCTGCGGCTGCGGCCCTTCGCCACGCGGAGCGCTTCACCGCTTCACGCCACCT 821
DB 266 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 325
QY 822 ATCCGTACTCGACACAGATCGACCTGCCACCCACCATCTCCTGTCTCAGCGGGAGG 881
DB 326 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 385
QY 882 AGCCCCACCTTACCAGGGGCGCTTGCACCTTCAGCTTCGGAGACCCCGAGCAGAGCTGG 941
DB 386 AGCCCCACCTTACCAGGGGCGCTTGCACCTTCAGCTTCGGAGACCCCGAGCAGAGCTGG 445

```

```

QY 942 AACTGAACGGGAGTGGTGGCGGCACACCCCAAAACAGAACCATCTTCGACAGTGCCTGA 1001
DB 446 AACTGAACGGGAGTGGTGGCGGCACACCCCAAAACAGAACCATCTTCGACAGTGCCTGA 505
QY 1002 TGGATAGTGCAGGCTGGCGGCCCTCCGCCCCAGCAGTAACCTCGGGCATCAGCGCA 1061
DB 506 TGGATAGTGCAGGCTGGCGGCCCTCCGCCCCAGCAGTAACCTCGGGCATCAGCGCA 565
QY 1062 COTGCTAGCGCAGCGCGGCGCGCATGAGAGGGGCGCGCCGCCCATCTACAGGAGTTCATCG 1121
DB 566 COTGCTAGCGCAGCGCGGCGCGCATGAGAGGGGCGCGCCGCCCATCTACAGGAGTTCATCG 625
QY 1122 GCCACTACCGGGGTCCTCTTCCAGCAGCAGCAGGAGTGGCGGCCCTCTCTGCTGG 1181
DB 626 GCCACTACCGGGGTCCTCTTCCAGCAGCAGCAGGAGTGGCGGCCCTCTCTGCTGG 685
QY 1182 AGGGACCGCGTCCACACACACATCGCGCCCTAGAGAGCGCAGCAGTCTGGAGCA 1241
DB 686 AGGGACCGCGTCCACACACACATCGCGCCCTAGAGAGCGCAGCAGTCTGGAGCA 745
QY 1242 AAGAGAGGATAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1276
DB 746 AAGAGAGGATAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
RESULT 13
AK008976
LOCUS
DEFINITION Mus musculus adult male stomach cDNA, RIKEN full-length enriched
library, clone:210418102 product:Need4 WW binding protein 4, full
insert sequence.
ACCESSION AK008976
VERSION AK008976.1 GI:12843488
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuana,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishi,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwara,S., Inoue,K., Izawa,M., Ogawa,Y., Okara,S., Watahiki,M.,
Yoneda,Y., Ishikawa,Y., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.

```


/note="Organ: Left Pubic Bone; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP ED1 is a normalized cDNA library containing the following tissue(s): Chondrosarcoma cell line C85. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CCTCAAGGCT.
TAG_TISSUE=chondrosarcoma
TAG_LIIB=UI-H-SD1
TAG_SEQ=CGTCAAGGCT"

ORIGIN

Query Match	46.5%	Score 614.6	DB 12	Length 782
Best Local Similarity	98.9%	Pred. No. 6.6e-76		
Matches 617	Conservative 0	Mismatches 7	Indels 0	Gaps 0
QY	698	GAGACACAGTGTACGCAACGGATCCAGAGCGCAGGTCTACGCCCGCCCTCGGCC	757	
Db	782	GAGACACAGTGTACGCAACGGATCCAGAGCGCAGGTNTACGCCCGCCCTCGGCC	723	
QY	758	ACCGACCGCTGGCGCTGGCGCCCTTCGCCAGCGGGAGCGCTTCCACCGCTTCAGGCC	817	
Db	722	ACCGACCGCTGGCGCTGGCGCCCTTCGCCAGCGGGAGCGCTTCCACCGCTTCAGGCC	663	
QY	818	ACCTATCCGTACTCTCAGCAGCAGATCGACCTGCCACCCACCATCTCGCTGTCTCAGACGG	877	
Db	662	ACCTATCCGTACTCTCAGCAGCAGATCGACCTGCCACCCACCATCTCGCTGTCTCAGACGG	603	
QY	878	GAGAGCCCCACCTTACAGGGCCCTTGACCCCTTCAGCTTCGGGACCCCGAGCAGCAG	937	
Db	602	GAGAGCCCCACCTTACAGGGCCCTTGACCCCTTCAGCTTCGGGACCCCGAGCAGCAG	543	
QY	938	CTGGAACTGAACCGGGAGTCGGTGGCGGCACCCGCCCAACAGAACCATCTTCGACAGTGAC	997	
Db	542	CTGGAACTGAACCGGGAGTCGGTGGCGGCACCCGCCCAACAGAACCATCTTCGACAGTGAC	483	
QY	998	CTGATGGATAGTGCACAGGTGGGGCGCCCTTCCCGCCCGCAGAGTAACCTCGGGCATCAGC	1057	
Db	482	CTGATGGATAGTGCACAGGTGGGGCGCCCTTCCCGCCCGCAGAGTAACCTCGGGCATCAGC	423	
QY	1058	GCCACGTCTACGCGAGCGCGCGCATGAGAGGGCGCGCCGCCACCTACAGCGAGGTC	1117	
Db	422	GCCACGTCTACGCGAGCGCGCGCATGAGAGGGCGCGCGCCGCCACCTACAGCGAGGTC	363	
QY	1118	ATCGGCCACTACCCGGGGTCTCTTCCAGCACACAGAGAGAGAGTGGGCGCCCTTCCTTG	1177	
Db	362	ATCGGCCACTACCCGGGGTCTCTTCCAGCACACAGAGAGAGAGTGGGCGCCCTTCCTTG	303	
QY	1178	CTGGAGGGACCCGGTCCACACACACATCGCGCCCTTAGAGAGCGCAGCATCTGG	1237	
Db	302	CTGGAGGGACCCGGTCCACACACACATCGCGCCCTTAGAGAGCGCAGCATCTGG	243	
QY	1238	AGCAAGAGAGAGATAAACAGAAAGACACCTCTCTAGGGTCCCCAGGGGGCGGGCT	1297	
Db	242	AGCAAGAGAGAGATAAACAGAAAGACACCTCTCTAGGGTCCCCAGGGGGCGGGCT	183	
QY	1298	GGGGTGCCTAGGTGAAAAGGCAG	1321	
Db	182	GGGGTGCCTAGGTGAAAAGGCAG	159	

Search completed: May 12, 2004, 05:18:05
Job time : 3988.63 sec